

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2003, 15:04:34 ; Search time 73 seconds

(without alignments)
399.752 Million cell updates/sec

Title: US-09-923-236-2

Perfect score: 219

Sequence: 1 MKLLWACIVFAFKRRF.....TAKPAPEPHPSLEQANQ 219

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 908470 seqs, 133250620 residues

Word size : 0

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A_Geneseq_101002.*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	219	100.0	219	22 AAB60109	Human transport pr
2	219	100.0	219	23 AAU74536	Human zsig63 polyp
3	150	68.5	221	20 AAY19472	Amino acid sequenc
4	140	63.9	219	20 AAW30653	Human secreted pro
5	9	4.1	75	22 ABB64202	Drosophila melanog
6	9	4.1	322	17 AAW03688	Leishmania chagasi
7	9	4.1	728	22 ABB65430	Drosophila melanog
8	9	4.1	1874	22 AAB76532	Corynebacterium gl
9	9	4.1	2969	22 AAG90880	C glutamicum prote
10	8	3.7	65	21 AAG27598	Arabidopsis thalia

11	8	3.7	82	21 AAG19327	Arabidopsis thalia
12	8	3.7	94	21 AAG19326	Arabidopsis thalia
13	8	3.7	106	21 AAG19325	Arabidopsis thalia
14	8	3.7	121	21 AAB40893	Human ORFX ORF657
15	8	3.7	124	21 AAG27837	Arabidopsis thalia
16	8	3.7	129	22 AAG63380	Amino acid sequenc
17	8	3.7	132	20 AAY34678	Chlamydia pneumoni
18	8	3.7	139	21 AAG54857	Arabidopsis thalia
19	8	3.7	145	21 AAB32749	Eucalyptus grandis
20	8	3.7	145	21 AAG07525	Arabidopsis thalia
21	8	3.7	145	21 AAG45122	Arabidopsis thalia
22	8	3.7	145	21 AAG45149	Arabidopsis thalia
23	8	3.7	181	21 AAG06866	Arabidopsis thalia
24	8	3.7	181	21 AAG10820	Arabidopsis thalia
25	8	3.7	185	21 AAG54822	Arabidopsis thalia
26	8	3.7	193	21 AAG06865	Arabidopsis thalia
27	8	3.7	193	21 AAG10819	Arabidopsis thalia
28	8	3.7	205	21 AAG06864	Arabidopsis thalia
29	8	3.7	205	21 AAG10818	Arabidopsis thalia
30	8	3.7	328	21 AAY90267	Protein chimera Q
31	8	3.7	372	22 AAG16669	Novel human diagno
32	8	3.7	412	21 AAY90266	Protein chimera Q
33	8	3.7	421	22 ABB67110	Drosophila melanog
34	8	3.7	428	23 AAM50963	Maize methyl Cpg 4b
35	8	3.7	493	22 ABB70327	Drosophila melanog
36	8	3.7	511	22 ABB28667	Peptide #1318 enco
37	8	3.7	511	22 ABB33852	Peptide #1358 enco
38	8	3.7	511	22 ABB19292	Protein #1291 enco
39	8	3.7	511	22 AAM54616	Human brain expres
40	8	3.7	511	22 AAM67022	Human bone marrow
41	8	3.7	511	22 AAM14883	Peptide #1317 enco
42	8	3.7	511	22 AAM27312	Peptide #1349 enco
43	8	3.7	511	22 AAM02607	Peptide #1289 enco
44	8	3.7	511	23 AAG36679	Human peptide enco
45	8	3.7	512	22 ABB61369	Drosophila melanog

ALIGNMENTS

RESULT 1

AAB60109
ID AAB60109 standard; Protein; 219 AA.

AC AC

XX XX

DT 28-MAR-2001 (first entry)

XX XX

DE Human transport protein TPPT-29.

XX XX

XX XX

KW Human; transport protein; TPPT; transport disorder; metabolic disorder;

KW neurological disorder; cardiovascular disorder; reproductive disorder;

XX XX

OS Homo sapiens.

XX XX

PN WO200078953-A2.

XX XX

PD 28-DEC-2000.

XX XX

PF 16-JUN-2000; 2000WO-US16668.

XX XX

PR 17-JUN-1999; 99US-0139923.

PR 10-AUG-1999; 99US-0148177.

PR 18-AUG-1999; 99US-0149357.

PR 28-OCT-1999; 99US-0162287.

XX XX

PA (INCY-) INCYTE GENOMICS INC.

XX XX

PI Lal P, Yang J, Yue H, Hillman JL, Tang YT, Bandman O, Burford N;

PI Baughn MR, Asimzai Y, Lu DAM, Au-Young J, Patterson C;

XX XX

DR WPI; 2001-041424/05.

DR N-PSDB; AAF27729.
XX Isolated polypeptide with a human transport protein sequence is useful
PT for the diagnosis, prevention and treatment of disorders associated
PT with the immune, reproductive and cardiovascular systems -
XX
XX Claim 2; Page 130; 165pp; English.
PS
CC The present invention provides the protein and coding sequences for 43
CC novel human transport proteins (designated TPTs). These can be used in
CC the diagnosis and treatment of transport, metabolic, neurological,
CC reproductive, cardiovascular and immune disorders, and cell proliferative
CC disorders such as cancer.
XX
XX Sequence 219 AA;
SQ

Query Match 100.0%; Score 219; DB 22; Length 219;
Best Local Similarity 100.0%; Pred. No. 8.3e-201;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKLLWACIVCVAFARKRPFPIGDDNDGHPHLSNIPYGINLPPLPYRPVNTVP 60
DB 1 MKLLWACIVCVAFARKRPFPIGDDNDGHPHLSNIPYGINLPPLPYRPVNTVP 60
QY 61 SYPGNTYTDGLSPYMWLTSPGFPYVYHIRGFLATQLNVPLPRGFPVPPSRFFSA 120
DB 61 SYPGNTYTDGLSPYMWLTSPGFPYVYHIRGFLATQLNVPLPRGFPVPPSRFFSA 120
QY 121 AAAPAPPIAAEPAAPLTPVAAPAGAPVAAEPAEAPVGAEPAAEAPVAAEPAA 180
DB 121 AAAPAPPIAAEPAAPLTPVAAPAGAPVAAEPAEAPVGAEPAAEAPVAAEPAA 180
QY 181 EAPVGEPAEAPSPAEPAATAKPAPEPHSPSLEQANQ 219
DB 181 EAPVGEPAEAPSPAEPAATAKPAPEPHSPSLEQANQ 219

RESULT 2
AAU74536
ID AAU74536 standard; Protein; 219 AA.

XX AC AAU74536;
XX
DT 23-APR-2002 (first entry)
XX
DE Human zsig63 polypeptide.
XX
KW Human: zsig63; chromosome 4q12-q13; salivary protein; antimicrobial;
KW microbial infection; tooth decay; periodontal disease; thrush; emphysema;
KW gastrointestinal disease; urinary tract infection; vaginal infection;
KW skin infection; epithelial wound; chronic tissue damage; cystic fibrosis;
KW acquired immunodeficiency syndrome; AIDS; lung infection; sarcoidosis;
KW chronic bronchitis; gene therapy; protein therapy.
XX
OS Homo sapiens.
XX
XX US6331413-B1.
XX
XX 18-DEC-2001.
XX
XX 17-MAR-2000; 2000US-0527345.
XX
XX 17-MAR-1999; 99US-124820P.
XX
XX (ZYMO) ZYMOGENETICS INC.
XX
XX Adler DA, Sheppard PO;
XX
XX WPI; 2002-096707/13.
DR N-PSDB; AAS20591.
XX
XX Polynucleotides encoding salivary proteins useful as anti-microbial
PT agents -

XX
PS Claim 1; Column 49-52; 29pp; English.
XX
CC The invention relates to a polynucleotide derived from the 4q12-q13
CC region of human chromosome 4 and encoding a zsig63 polypeptide, a
CC secreted salivary protein with anti-microbial activity. Due to their
CC microbial activity, the sequences can be used in the study of microbial
CC infections, e.g. for recombinant production of anti-microbial proteins.
CC The sequences can be used in the treatment of tooth decay, periodontal
CC disease, thrush, gastrointestinal disease, urinary tract infections,
CC vaginal infections, skin infections, epithelial wounds, chronic tissue
CC damage, acquired immunodeficiency syndrome (AIDS), cystic fibrosis, lung
CC infections, sarcoidosis, emphysema and chronic bronchitis. This sequence
CC represents human zsig63.
XX
XX Sequence 219 AA;
SQ

Query Match 100.0%; Score 219; DB 23; Length 219;
Best Local Similarity 100.0%; Pred. No. 8.3e-201;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKLLWACIVCVAFARKRPFPIGDDNDGHPHLSNIPYGINLPPLPYRPVNTVP 60
DB 1 MKLLWACIVCVAFARKRPFPIGDDNDGHPHLSNIPYGINLPPLPYRPVNTVP 60
QY 61 SYPGNTYTDGLSPYMWLTSPGFPYVYHIRGFLATQLNVPLPRGFPVPPSRFFSA 120
DB 61 SYPGNTYTDGLSPYMWLTSPGFPYVYHIRGFLATQLNVPLPRGFPVPPSRFFSA 120
QY 121 AAAPAPPIAAEPAAPLTPVAAPAGAPVAAEPAEAPVGAEPAAEAPVAAEPAA 180
DB 121 AAAPAPPIAAEPAAPLTPVAAPAGAPVAAEPAEAPVGAEPAAEAPVAAEPAA 180
QY 181 EAPVGEPAEAPSPAEPAATAKPAPEPHSPSLEQANQ 219
DB 181 EAPVGEPAEAPSPAEPAATAKPAPEPHSPSLEQANQ 219

RESULT 3
AAU19472

ID AAU19472 standard; Protein; 221 AA.
XX
AC AAU19472;
XX
DT 14-JUL-1999 (first entry)
XX
DE Amino acid sequence of a human secreted protein.
XX
KW Human secreted protein; cancer; tumour; neurodegenerative disorder;
KW developmental abnormality; fetal deficiency; blood disorder; leukemia;
KW immune system disease; autoimmune disease; hepatic disease;
KW renal disease; lymphoma; inflammation; allergy; ischemic shock;
KW Alzheimer's; cognitive disorder; schizophrenia; prostate disease;
KW obesity; osteoclast; osteoporosis; arthritis; malignancy; testes disease;
KW lung disease; thymus disease; digestive disorder; endocrine disorder;
KW infection; AIDS.
XX
XX Homo sapiens.
XX
XX WO9922243-A1.
XX
XX 06-MAY-1999.
XX
XX 23-OCT-1998; 98WO-US22376.
XX
XX 24-OCT-1997; 97US-0063387.
PR 24-OCT-1997; 97US-0062784.
PR 24-OCT-1997; 97US-0063088.
PR 24-OCT-1997; 97US-0063089.
PR 24-OCT-1997; 97US-0063090.
PR 24-OCT-1997; 97US-0063091.
PR 24-OCT-1997; 97US-0063092.
PR 24-OCT-1997; 97US-0063097.

PR 24-OCT-1997; 97US-0063098.
 PR 24-OCT-1997; 97US-0063099.
 PR 24-OCT-1997; 97US-0063100.
 PR 24-OCT-1997; 97US-0063101.
 PR 24-OCT-1997; 97US-0063109.
 PR 24-OCT-1997; 97US-0063110.
 PR 24-OCT-1997; 97US-0063111.
 PR 24-OCT-1997; 97US-0063114.
 PR 24-OCT-1997; 97US-0063386.
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX
 PI Brewer LA, Carter KC, Duan DR, Ebner R, Endress CA;
 PI Feng P, Florence C, Florence KA, Greene JM, Janat F;
 PI Kay H, Lafleur DW, Moore PA, Ni J, Olsen HS, Rosen CA;
 PI Ruben SM, Shi Y, Soppet DR, Wei Y, Young P;
 XX
 DR WPI: 1999-303069/25.
 DR N-PSDB; AAX61352.
 XX
 XX
 PT New isolated human genes and the secreted polypeptides they encode
 XX
 XX
 PS Claim 11; Page 401-402; 546pp; English.
 XX
 CC The specification describes cDNA sequences (AAX61322-X61470) encoding
 CC human secreted proteins (AAY19442-Y19590). The polynucleotides and their
 CC corresponding secreted polypeptides are useful for preventing, treating
 CC or ameliorating medical conditions, e.g. by protein or gene therapy.
 CC Pathological conditions can also be diagnosed by determining the amount
 CC of the polypeptides in a sample or by determining the presence of
 CC mutations in the polynucleotides. Specific uses are described for each
 CC of the polynucleotides, based on which tissues they are most highly
 CC expressed in, and include developing products for the diagnosis or
 CC treatment of cancer, tumours, neurodegenerative disorders, developmental
 CC abnormalities and fetal deficiencies, blood disorders, leukemias,
 CC diseases of the immune system, autoimmune diseases, hepatic and renal
 CC disease, lymphomas, inflammation, allergies, ischemic shock, Alzheimer's
 CC and cognitive disorders, schizophrenia, prostate diseases, obesity,
 CC disorders involving osteoclasts such as osteoporosis, arthritis or
 CC malignancies, diseases of testes, lung or thymus, digestive/endocrine
 CC disorders, infections and AIDS. The polypeptides are also useful for
 CC identifying their binding partners.
 XX
 SQ Sequence 221 AA;
 Query Match 68.5%; Score 150; DB 20; Length 221;
 Best Local Similarity 100.0%; Pred. No. 5 6e-135;
 Matches 150; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MKLLWACIVCVAFARKRRPFIFGEDDDGHPHPSLNIPYGINLPPLPYRPNVTVP 60
 DB 1 MKLLWACIVCVAFARKRRPFIFGEDDDGHPHPSLNIPYGINLPPLPYRPNVTVP 60
 QY 61 SYPGNTYDTGLSPYHILSPGFPYVYHGRFPLATQLNVPPLPRGFFVPPSRFFSA 120
 DB 61 SYPGNTYDTGLSPYHILSPGFPYVYHGRFPLATQLNVPPLPRGFFVPPSRFFSA 120
 QY 121 AAAPAAPPIAAEPAAAPLTPVAEPAA 150
 DB 121 AAAPAAPPIAAEPAAAPLTPVAEPAA 150
 RESULT 4
 AAW30653
 ID AAW30653 standard; Protein; 219 AA.
 XX
 AC AAW30653;
 XX
 DT 12-APR-1999 (first entry)
 XX
 DE Human secreted protein clone cp116 1 protein.
 XX
 XX Human; secreted protein; nutritional activity; cytokine; vaccine;

KW cell proliferation; differentiation; immune stimulation; suppression;
 KW hematopoiesis regulation; tissue growth; activin; inhibin; chemotactic;
 KW chemokinetic; haemostatic; thrombolytic; anti-inflammatory; gene therapy;
 KW tumour invasion suppression; tumour inhibition.
 XX Homo sapiens.
 XX WO9901466-A1.
 PN 14-JAN-1999.
 PD 01-JUL-1998; 98WO-US13813.
 XX 27-OCT-1997; 97US-0958304.
 PR 02-JUL-1997; 97US-0887195.
 XX (GEMY) GENETICS INST INC.
 PA Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Racie LA;
 XX Spaulding V, Treacy M;
 PI WPI: 1999-105994/09.
 DR N-PSDB; AAV80740.
 XX
 XX New polynucleotides encoding secreted human proteins - are derived
 PT from human foetal brain, adult testes, adult brain, foetal kidney,
 PT adult salivary gland, or adult blood cDNA libraries, useful as, e.g.
 PT potential vaccines
 XX
 PS Claim 24; Page 71-72; 107pp; English.
 XX
 CC The present sequence represents a human secreted protein from clone
 CC cp116 1, deposited as ATCC 98482. Human secreted protein clone
 CC polynucleotides and proteins are predicted to have biological
 CC activities which would make them suitable for treating, preventing or
 CC ameliorating medical conditions in humans and animals. Suggested
 CC activities include nutritional activity, cytokine and cell
 CC proliferation/differentiation activity, immune stimulating (e.g. as
 CC vaccines) or suppressing activity, haematopoiesis regulating activity,
 CC tissue growth activity, activin/inhibin activity, chemotactic/
 CC chemokinetic activity, haemostatic and thrombolytic activity, receptor/
 CC ligand activity, anti-inflammatory activity, cadherin/tumour invasion
 CC suppressor activity, and tumour inhibition activity. The polynucleotides
 CC are also stated to be useful for gene therapy.
 XX
 SQ Sequence 219 AA;
 Query Match 63.9%; Score 140; DB 20; Length 219;
 Best Local Similarity 100.0%; Pred. No. 1.9e-125;
 Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MKLLWACIVCVAFARKRRPFIFGEDDDGHPHPSLNIPYGINLPPLPYRPNVTVP 60
 DB 1 MKLLWACIVCVAFARKRRPFIFGEDDDGHPHPSLNIPYGINLPPLPYRPNVTVP 60
 QY 61 SYPGNTYDTGLSPYHILSPGFPYVYHGRFPLATQLNVPPLPRGFFVPPSRFFSA 120
 DB 61 SYPGNTYDTGLSPYHILSPGFPYVYHGRFPLATQLNVPPLPRGFFVPPSRFFSA 120
 QY 121 AAAPAAPPIAAEPAAAPLTP 140
 DB 121 AAAPAAPPIAAEPAAAPLTP 140
 RESULT 5
 ABB64202
 ID ABB64202 standard; Protein; 75 AA.
 XX
 AC ABB64202;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 19398.

XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX Drosophila melanogaster.
OS WO200171042-A2.
XX 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US09231.
XX 23-MAR-2000; 2000US-191637P.
XX 11-JUL-2000; 2000US-0614150.
XX (PEKE) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW;
PI WPI: 2001-656860/75.
XX N-PSDB; ABL08305.
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX Disclosure; SEQ ID NO 19398; 21pp + Sequence Listing; English.
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 75 AA;
Query Match 4.1%; Score 9; DB 22; Length 75;
Best Local Similarity 100.0%; Pred. No. 0.76;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 120 AAAAPAAP 128
DB 45 AAAAPAAP 53
RESULT 6
AAW03688
ID AAW03688 standard; Protein; 322 AA.
XX
AC AAW03688;
XX 09-MAR-1997 (first entry)
XX Leishmania chagasi acidic ribosomal antigen LcP0.
XX Leishmania chagasi; acidic ribosomal antigen; LcP0;
KW epitope; K39.
XX Leishmania chagasi.
OS
XX Key Location/Qualifiers
FH CDS 30..1202
FT /*tag= a
XX WO9633414-A2.
XX 24-OCT-1996.

XX 19-APR-1996; 96WO-US05472.
XX 21-APR-1995; 95US-0428414.
XX (CORI-) CORIXA CORP.
XX Reed SG;
XX WPI: 1996-485884/48.
XX N-PSDB; AAT42164.
XX New Leishmania acidic ribosomal P-protein family poly:peptide - used
PT to develop prods. for diagnosis, detection and protection against
PT Leishmania infections
XX Disclosure; Page 29-32; 76pp; English.
XX Compounds including polypeptides that contain at least an epitope of
CC the L. chagasi acidic ribosomal antigen LcP0 are useful in a variety
CC of immunoassays for detecting Leishmania infection. Portions of
CC LcP0 (AAT42164) contg. at least the 17 C-terminal amino acids (AAT42165)
CC have been found to generate a signal in an ELISA that is equivalent
CC to that generated by the full length LcP0. A combination
CC polypeptide may also be used, comprising an LcP0 epitope along with
CC an epitope derived from the Leishmania K39 antigen (AAT42166), pref.
CC the K39 repeat unit antigen having the sequence given in AAW03690.
XX
SQ Sequence 322 AA;
Query Match 4.1%; Score 9; DB 17; Length 322;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 130 AAEPAAAP 138
DB 289 AAEPAAAP 297
RESULT 7
ABB65430
ID ABB65430 standard; Protein; 728 AA.
XX
AC ABB65430;
XX 26-MAR-2002 (first entry)
XX Drosophila melanogaster polypeptide SEQ ID NO 23082.
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX Drosophila melanogaster.
OS
XX WO200171042-A2.
XX 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US09231.
XX 23-MAR-2000; 2000US-191637P.
XX 11-JUL-2000; 2000US-0614150.
XX (PEKE) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW;
PI WPI: 2001-656860/75.
XX N-PSDB; ABL09533.
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -

XX Disclosure; SEQ ID NO 23082; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent

XX capable of detecting 1000 or more genes from *Drosophila*. The invention is

CC useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

CC sequences (ABL01840-ABL16175) and the encoded proteins

CC (ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX

XX Sequence 728 AA;

SQ

Query Match 4.1%; Score 9; DB 22; Length 728;

Best Local Similarity 100.0%; Pred. No. 5.1;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 119 SAARAPAP 127

DB 418 SAARAPAP 426

|||||||

RESULT 8

AAAB76532

ID AAB76532 standard; Protein; 1874 AA.

XX

AC AAB76532;

XX

DT 11-APR-2001 (first entry)

XX

DE Corynebacterium glutamicum MCT protein SEQ ID NO:46.

XX

XX Corynebacterium glutamicum; brevibacterium lactofermentum; MCT;

KW membrane construction and membrane transport protein; petroleum spill;

KW hydrocarbon degradation; gram positive aerobic bacterium; marker;

KW identification; microorganism; fine chemical production; transformation;

KW genome mapping; genetic engineering.

XX

OS Corynebacterium glutamicum.

XX

PN WO200100805-A2.

XX

PD 04-JAN-2001.

XX

PF 23-JUN-2000; 2000WO-IB00926.

XX

PR 25-JUN-1999; 99US-0141031.

PR 08-JUL-1999; 99DE-1031454.

PR 08-JUL-1999; 99DE-1031478.

PR 08-JUL-1999; 99DE-1031563.

PR 09-JUL-1999; 99DE-1032122.

PR 09-JUL-1999; 99DE-1032124.

PR 09-JUL-1999; 99DE-1032125.

PR 09-JUL-1999; 99DE-1032128.

PR 09-JUL-1999; 99DE-1032180.

PR 09-JUL-1999; 99DE-1032182.

PR 09-JUL-1999; 99DE-1032190.

PR 09-JUL-1999; 99DE-1032191.

PR 09-JUL-1999; 99DE-1032209.

PR 09-JUL-1999; 99DE-1032212.

PR 09-JUL-1999; 99DE-1032227.

PR 09-JUL-1999; 99DE-1032228.

PR 09-JUL-1999; 99DE-1032229.

PR 09-JUL-1999; 99DE-1032230.

PR 14-JUL-1999; 99DE-1032927.

PR 14-JUL-1999; 99DE-1033005.

PR 14-JUL-1999; 99DE-1033006.

PR 27-AUG-1999; 99DE-1040764.

PR 27-AUG-1999; 99DE-1040765.

PR 27-AUG-1999; 99DE-1040766.

PR 27-AUG-1999; 99DE-1040830.

PR 27-AUG-1999; 99DE-1040831.

PR 27-AUG-1999; 99DE-1040832.

PR 27-AUG-1999; 99DE-1040833.

PR 31-AUG-1999; 99DE-1041378.

PR 31-AUG-1999; 99DE-1041379.

PR 31-AUG-1999; 99DE-1041395.

PR 03-SEP-1999; 99DE-1042077.

PR 03-SEP-1999; 99DE-1042078.

PR 03-SEP-1999; 99DE-1042079.

PR 03-SEP-1999; 99DE-1042088.

XX

PA (BADI) BASF AG.

XX

PI Pompejus M., Kroeger B, Schroeder H, Zelder O, Haberhauer G;

XX

DR WPI; 2001-071486/08.

DR N-PSDB; AAF67765.

XX

PT Corynebacterium glutamicum nucleic acids encoding membrane construction

PT and membrane transport proteins or their portions, useful for typing or

PT identifying C. glutamicum or related bacteria, and as markers for

PT transformation.

XX

PS Claim 20; Page 202-207; 1119pp; English.

XX

CC AAF67743 to AAF68080 encode the Corynebacterium glutamicum membrane

CC construction and membrane transport (MCT) proteins given in AAB76510 to

CC AAB76847. The MCT nucleic acids and proteins are useful in the

CC identification of microorganisms which can be used to produce fine

CC chemicals, for modulating fine chemical production in C. glutamicum or

CC related bacteria (e.g. Brevibacterium lactofermentum), the typing or

CC identification of C. glutamicum or related bacteria, as reference points

CC for mapping C. glutamicum genome, and as markers for transformation.

CC AAF68082 and AAF68082 represent sequencing primers which are used in an

CC example from the present invention.

XX

SQ Sequence 1874 AA;

Query Match 4.1%; Score 9; DB 22; Length 1874;

Best Local Similarity 100.0%; Pred. No. 11;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 168 PRAEAPVAA 176

DB 568 PRAEAPVAA 576

|||||||

RESULT 9

AAG90680

ID AAG90680 standard; Protein; 2969 AA.

XX

AC AAG90680;

XX

DT 26-SEP-2001 (first entry)

XX

DE C glutamicum protein fragment SEQ ID NO: 4434.

XX

KW Coryneform bacterium; amino acid synthesis; vitamin; saccharide;

KW organic acid synthesis.

XX

OS Corynebacterium glutamicum.

XX

PN EF1108790-A2.

XX

PD 20-JUN-2001.

XX

PF 18-DEC-2000; 2000EP-0127688.

XX

PR 16-DEC-1999; 99JP-0377484.

PR 07-APR-2000; 2000JP-0159162.

PR 03-AUG-2000; 2000JP-0280988.

```

XX (KYOW ) KYOWA HAKKO KOGYO KK.
PA
XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
PI
XX WPI: 2001-376931/40.
DR N-PSDB; AAH65899.
DR
XX Novel polynucleotides derived from Coryneform bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analysing
PT expression profile or pattern of a gene and identifying homologous gene
PT
XX
PS Claim 17; SEQ ID NO: 4434; 246pp + Sequence Listing; English.
PS
XX The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and
CC analysing the expression profile or expression pattern of a gene derived
CC from Coryneform bacterium, and identifying a homologue of a gene derived
CC from coryneform bacterium. Coryneform bacteria are useful for producing
CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a protein described
CC in the exemplification of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office.
XX
XX Sequence 2969 AA;
SQ
Query Match 4.1%; Score 9; DB 22; Length 2969;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 168 PAEAPVAA 176
Db 1663 PAEAPVAA 1671
RESULT 10
AAG27598
ID AAG27598 standard; Protein; 65 AA.
XX
AC AAG27598;
XX
XX 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 32499.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
99US-0130077.
99US-0130449.
99US-0130510.
99US-0130891.
99US-0131449.
99US-0132048.
99US-0132407.
99US-0132484.
99US-0132485.
99US-0132486.
99US-0132487.
99US-0132863.
99US-0132863.
99US-0134256.
99US-0134218.
99US-0134219.
99US-0134221.
99US-0134370.
99US-0134768.
99US-0134941.
99US-0135124.
99US-0135353.
99US-0135629.
99US-0136021.
99US-0136392.
99US-0136782.
99US-0137222.
99US-0137528.
99US-0137502.
99US-0137724.
99US-0138094.
99US-0138540.
99US-0138647.
99US-0139119.
99US-0139452.
99US-0139453.
99US-0139492.
99US-0139454.
99US-0139455.
99US-0139456.
99US-0139457.
99US-0139458.
99US-0139459.
99US-0139460.
99US-0139461.
99US-0139462.
99US-0139463.
99US-0139750.
99US-0139763.
99US-0139817.
99US-0139899.
99US-0140353.
99US-0140354.
99US-0140695.
99US-0140823.
99US-0140991.
99US-0141287.
99US-0141842.
99US-0142154.
99US-0142055.
99US-0142390.
99US-0142803.
99US-0142920.
99US-0142977.
99US-0143542.
99US-0143624.
99US-0144005.
99US-0144085.
99US-0144086.
99US-0144325.
99US-0144331.
99US-0144332.
99US-0144333.
99US-0144334.
99US-0144334.

```

```
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145313.
PR 27-JUL-1999; 99US-0145318.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 12-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.

PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 3.7%; Score 8; DB 21; Length 65;
Best Local Similarity 100.0%; Pred. No. 6.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 AAEPAAAA 137
DB 19 AAEPAAAA 26
|||||||

RESULT 11
AAG19327
ID AAG19327 standard; Protein; 82 AA.
XX AC AAG19327;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 21085.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX OS Arabidopsis thaliana.
XX PN EPI033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PP 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
```

PR	04-MAY-1999;	99US-0132484.	PR	22-JUL-1999;	99US-0145085
PR	05-MAY-1999;	99US-0132485.	PR	22-JUL-1999;	99US-0145087
PR	06-MAY-1999;	99US-0132486.	PR	22-JUL-1999;	99US-0145089
PR	06-MAY-1999;	99US-0132487.	PR	22-JUL-1999;	99US-0145192
PR	07-MAY-1999;	99US-0132863.	PR	23-JUL-1999;	99US-0145145.
PR	11-MAY-1999;	99US-0134256.	PR	23-JUL-1999;	99US-0145218.
PR	14-MAY-1999;	99US-0134218.	PR	23-JUL-1999;	99US-0145224.
PR	14-MAY-1999;	99US-0134219.	PR	26-JUL-1999;	99US-0145276.
PR	14-MAY-1999;	99US-0134221.	PR	27-JUL-1999;	99US-0145913.
PR	14-MAY-1999;	99US-0134370.	PR	27-JUL-1999;	99US-0145918.
PR	18-MAY-1999;	99US-01343768.	PR	27-JUL-1999;	99US-0145919.
PR	19-MAY-1999;	99US-0134941.	PR	28-JUL-1999;	99US-0145951.
PR	20-MAY-1999;	99US-0135124.	PR	02-AUG-1999;	99US-0146388.
PR	21-MAY-1999;	99US-0135353.	PR	02-AUG-1999;	99US-0146389.
PR	24-MAY-1999;	99US-0135629.	PR	02-AUG-1999;	99US-0147038.
PR	25-MAY-1999;	99US-0136021.	PR	03-AUG-1999;	99US-0147204.
PR	27-MAY-1999;	99US-0136392.	PR	04-AUG-1999;	99US-0147302.
PR	28-MAY-1999;	99US-0136782.	PR	04-AUG-1999;	99US-0147302.
PR	01-JUN-1999;	99US-0137222.	PR	05-AUG-1999;	99US-0147260.
PR	03-JUN-1999;	99US-0137528.	PR	05-AUG-1999;	99US-0147303.
PR	04-JUN-1999;	99US-0137502.	PR	06-AUG-1999;	99US-0147416.
PR	07-JUN-1999;	99US-0137724.	PR	06-AUG-1999;	99US-0147493.
PR	08-JUN-1999;	99US-0138094.	PR	09-AUG-1999;	99US-0147935.
PR	10-JUN-1999;	99US-0138540.	PR	10-AUG-1999;	99US-0148171.
PR	10-JUN-1999;	99US-0138847.	PR	11-AUG-1999;	99US-0148319.
PR	14-JUN-1999;	99US-0139119.	PR	12-AUG-1999;	99US-0148341.
PR	16-JUN-1999;	99US-0139452.	PR	13-AUG-1999;	99US-0148565.
PR	16-JUN-1999;	99US-0139453.	PR	13-AUG-1999;	99US-0148684.
PR	17-JUN-1999;	99US-0139492.	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	99US-0139455.	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	99US-0139456.	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	99US-0139457.	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	99US-0139458.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139459.	PR	20-AUG-1999;	99US-0149729.
PR	18-JUN-1999;	99US-0139460.	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	99US-0139461.	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	99US-0139462.	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	99US-0139463.	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	99US-0139465.	PR	27-AUG-1999;	99US-0151065.
PR	18-JUN-1999;	99US-0139750.	PR	27-AUG-1999;	99US-0151066.
PR	18-JUN-1999;	99US-0139763.	PR	27-AUG-1999;	99US-0151080.
PR	21-JUN-1999;	99US-0139817.	PR	30-AUG-1999;	99US-0151303.
PR	22-JUN-1999;	99US-0139899.	PR	31-AUG-1999;	99US-0151438.
PR	23-JUN-1999;	99US-0140353.	PR	01-SEP-1999;	99US-0151930.
PR	23-JUN-1999;	99US-0140354.	PR	01-SEP-1999;	99US-0152363.
PR	24-JUN-1999;	99US-0140695.	PR	07-SEP-1999;	99US-0153070.
PR	24-JUN-1999;	99US-0140823.	PR	10-SEP-1999;	99US-0153758.
PR	28-JUN-1999;	99US-0140931.	PR	13-SEP-1999;	99US-0154018.
PR	29-JUN-1999;	99US-0141287.	PR	15-SEP-1999;	99US-0154039.
PR	30-JUN-1999;	99US-0141842.	PR	16-SEP-1999;	99US-0154079.
PR	01-JUL-1999;	99US-01421154.	PR	20-SEP-1999;	99US-0155139.
PR	02-JUL-1999;	99US-0142055.	PR	22-SEP-1999;	99US-0155139.
PR	06-JUL-1999;	99US-0142390.	PR	23-SEP-1999;	99US-0155486.
PR	08-JUL-1999;	99US-0142803.	PR	24-SEP-1999;	99US-


```
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160988.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match      3.7%; Score 8; DB 21; Length 82;
Best Local Similarity 100.0%; Pred. No. 7.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 AAAAPAP 127
Db      |||||
      8 AAAAPAP 15

RESULT 12
AAG19326
ID AAG19326 standard; Protein; 94 AA.
XX AC
XX AAG19326;
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 21084.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144634.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
```

```
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148584.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149829.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 23-SEP-1999; 99US-0155659.
PR 24-SEP-1999; 99US-0155658.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0160993.

PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

      Query Match      3.7%; Score 8; DB 21; Length 94;
      Best Local Similarity 100.0%; Pred.No. 8.2;
      Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 AAAPAAP 127
DB 20 AAAPAAP 27
|||||||

RESULT 13
AAG19325
ID AAG19325 standard; Protein; 106 AA.
XX
AC AAG19325;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 21083.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
```

PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137228.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139111.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145099.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147835.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 13-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.

PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 3.7%; Score 8; DB 21; Length 106;
Best Local Similarity 100.0%; Pred. No. 9.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 120 AAAAPAP 127
Db 32 AAAAPAP 39
|||||||

RESULT 14
AAB40893
ID AAB40893 standard; Protein; 121 AA.
XX AC AAB40893;
XX DT
XX 08-FEB-2001 (first entry)
XX DE Human ORFX ORF657 polypeptide sequence SEQ ID NO:1314.
XX
XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive.
XX OS Homo sapiens.
XX PN W0200058473-A2.
XX XX
XX 05-OCT-2000.
XX PF
XX 31-MAR-2000; 2000WO-US08621.
XX PR 31-MAR-1999; 99US-0127607.
XX PR 02-APR-1999; 99US-0127636.
XX PR 05-APR-1999; 99US-0127728.
XX PR 30-MAR-2000; 2000US-0540763.
XX XX
XX (CURA-) CURAGEN CORP.
XX XX
XX Shinkets RA, Leach M;
PI
XX WPI; 2000-602362/57.
DR N-PSDB; AAC75102.
XX
XX Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease -
XX
XX Claim 11; Page 1131; 5507pp; English.
XX
XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnary;
CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC antidiabetic; hypotensive; dermatological; immunosuppressive;
CC antiinflammatory; antibacterial; antiviral; antirheumatic;
CC antithyroid; and antianaemic. The sequences can be used for determining

CC the presence of or predisposition to, or preventing or treating
CC pathological conditions associated with an ORFX-associated disorder. The
CC nucleic acids can be used to express ORFX proteins in gene therapy
CC vectors. The proteins and nucleic acids may be used to treat cancers,
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC graft vs host disease, cardiovascular disease, diabetes mellitus,
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
CC coagulation; to inhibit thrombosis; and as a contraceptive.
XX
SQ Sequence 121 AA;

Query Match 3.7%; Score 8; DB 21; Length 121;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 121 AAAAPAPP 128
Db 103 AAAAPAPP 110
|||||||

RESULT 15
AAG27837
ID AAG27837 standard; Protein; 124 AA.
XX AC AAG27837;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 32830.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX XX
XX 06-SEP-2000.
XX PF
XX 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
XX PR 05-MAR-1999; 99US-0123180.
XX PR 09-MAR-1999; 99US-0123548.
XX PR 23-MAR-1999; 99US-0125788.
XX PR 25-MAR-1999; 99US-0126264.
XX PR 29-MAR-1999; 99US-0126785.
XX PR 01-APR-1999; 99US-0127462.
XX PR 06-APR-1999; 99US-0128234.
XX PR 08-APR-1999; 99US-0128714.
XX PR 16-APR-1999; 99US-0129845.
XX PR 19-APR-1999; 99US-0130077.
XX PR 21-APR-1999; 99US-0130449.
XX PR 23-APR-1999; 99US-0130510.
XX PR 28-APR-1999; 99US-0130891.
XX PR 30-APR-1999; 99US-0131449.
XX PR 30-APR-1999; 99US-0132048.
XX PR 30-APR-1999; 99US-0132407.
XX PR 04-MAY-1999; 99US-0132484.
XX PR 05-MAY-1999; 99US-0132485.
XX PR 06-MAY-1999; 99US-0132486.
XX PR 07-MAY-1999; 99US-0132863.
XX PR 11-MAY-1999; 99US-0134256.
XX PR 14-MAY-1999; 99US-0134218.
XX PR 14-MAY-1999; 99US-0134219.
XX PR 14-MAY-1999; 99US-0134221.
XX PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137722.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140699.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.

PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161922.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 3.7%; Score 8; DB 21; Length 124;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 130 AAEPA AAA 137
 |||||
 Db 19 AAEPA AAA 26

Search completed: July 7, 2003, 15:10:35
 Job time : 77 secs

NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/428,414A
FILING DATE: 21-APR-1995
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Kadlecěk, Ann T.
REGISTRATION NUMBER: 39,244
REFERENCE/DOCKET NUMBER: 210121.407
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
TELEX: 3723836 SEEDANDBERRY
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 322 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-428-414A-2

Query Match 4.1%; Score 9; DB 2; Length 322;
Best Local Similarity 100.0%; Pred. No. 0.57;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 AAEPAAAP 138
DB 289 AAEPAAAP 297

RESULT 3
US-09-471-396-3
; Sequence 3, Application US/09471396
; Patent No. 6458359
; GENERAL INFORMATION:
; APPLICANT: BEDATE, Carlos Alonso
; APPLICANT: REQUENA ROLANIA, Jose Maria
; TITLE OF INVENTION: CHIMERIC GENE FORMED OF THE DNA SEQUENCES THAT ENCODE
; TITLE OF INVENTION: THE ANTIGENIC DETERMINANTS OF FOUR PROTEINS OF
; TITLE OF INVENTION: L. INFANTUM AND PROTEIN ENCODED BY SAIDI GENE, AND ...
; FILE REFERENCE: bedate2a.seq
; CURRENT APPLICATION NUMBER: US/09/471,396
; CURRENT FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: 60/113,825
; PRIOR FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Protein Q
US-09-471-396-3

Query Match 3.7%; Score 8; DB 4; Length 383;
Best Local Similarity 100.0%; Pred. No. 6.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 AAAPAAAP 127

DB 354 AAAAPAAP 361

RESULT 4
US-09-471-396-1
; Sequence 1, Application US/09471396
; Patent No. 6458359
; GENERAL INFORMATION:
; APPLICANT: BEDATE, Carlos Alonso
; APPLICANT: REQUENA ROLANIA, Jose Maria
; TITLE OF INVENTION: CHIMERIC GENE FORMED OF THE DNA SEQUENCES THAT ENCODE
; TITLE OF INVENTION: THE ANTIGENIC DETERMINANTS OF FOUR PROTEINS OF
; TITLE OF INVENTION: L. INFANTUM AND PROTEIN ENCODED BY SAIDI GENE, AND ...
; FILE REFERENCE: bedate2a.seq
; CURRENT APPLICATION NUMBER: US/09/471,396
; CURRENT FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: 60/113,825
; PRIOR FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Protein Q
US-09-471-396-1

Query Match 3.7%; Score 8; DB 4; Length 412;
Best Local Similarity 100.0%; Pred. No. 6.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 AAAAPAAP 127
DB 383 AAAAPAAP 390

RESULT 5
US-08-653-648A-5
; Sequence 5, Application US/08653648A
; Patent No. 6379945
; GENERAL INFORMATION:
; APPLICANT: Jepson, Ian
; APPLICANT: Greenland, Andrew
; APPLICANT: Martinez, Alberto
; TITLE OF INVENTION: A Gene Switch
; FILE REFERENCE: PPD50047/US
; CURRENT APPLICATION NUMBER: US/08/653,648A
; CURRENT FILING DATE: 1996-05-24
; PRIOR APPLICATION NUMBER: GB 9510759.5
; PRIOR FILING DATE: 1995-05-26
; PRIOR APPLICATION NUMBER: GB 9605656.9
; PRIOR FILING DATE: 1996-03-18
; PRIOR APPLICATION NUMBER: GB 9513882.2
; PRIOR FILING DATE: 1995-07-07
; PRIOR APPLICATION NUMBER: GB 9517316.7
; PRIOR FILING DATE: 1995-08-24
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 575
; TYPE: PRT
; ORGANISM: Heliothis virescens
US-08-653-648A-5

Query Match 3.7%; Score 8; DB 4; Length 575;
Best Local Similarity 100.0%; Pred. No. 9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 140 TATPVAEE 147

Db 552 TATPVAE 559

RESULT 6

US-08-804-227C-14
; Sequence 14, Application US/08804227C
; Patent No. 5876991
; GENERAL INFORMATION:
; APPLICANT: DeHoff, Bradley S.
; APPLICANT: Kustoss, Stuart A.
; APPLICANT: Rostock, Paul R., Jr.
; APPLICANT: Sutton, Kimberly L.
; TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: THOMAS G. PLANT 1501
; STREET: LILLY CORPORATE CENTER
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: ASCII(DOS) Text only
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,227C
; FILING DATE: February 21, 1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Plant, Thomas, G.
; REGISTRATION NUMBER: 35,784
; REFERENCE/DOCKET NUMBER: X-8231
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-2459
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4545 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-804-227C-14

Query Match 3.7%; Score 8; DB 2; Length 4545;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 AEPAAAAAP 138

Db 2620 AEPAAAAAP 2627

RESULT 7

US-08-804-227C-8
; Sequence 8, Application US/08804227C
; Patent No. 5876991
; GENERAL INFORMATION:
; APPLICANT: DeHoff, Bradley S.
; APPLICANT: Kustoss, Stuart A.
; APPLICANT: Rostock, Paul R., Jr.
; APPLICANT: Sutton, Kimberly L.
; TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: THOMAS G. PLANT 1501
; STREET: LILLY CORPORATE CENTER
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: ASCII(DOS) Text only
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,227C
; FILING DATE: February 21, 1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Plant, Thomas, G.
; REGISTRATION NUMBER: 35,784
; REFERENCE/DOCKET NUMBER: X-8231
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-2459
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4550 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-804-227C-8

Query Match 3.7%; Score 8; DB 2; Length 4550;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 AEPAAAAAP 138

Db 2625 AEPAAAAAP 2632

RESULT 8

US-08-804-198-2
; Sequence 2, Application US/08804198
; Patent No. 5945320
; GENERAL INFORMATION:
; APPLICANT: Burgett, Stanley G.
; APPLICANT: Kustoss, Stuart A.
; APPLICANT: Rao, Nagaraja R.
; APPLICANT: Richardson, Mark A.
; APPLICANT: Rostock, Paul R., Jr.
; TITLE OF INVENTION: PLATENOLIDE SYNTHASE GENE
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PAUL R. CANTRELL 1138
; STREET: LILLY CORPORATE CENTER
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: Macintosh 7.0
; SOFTWARE: Microsoft Word 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,198
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: CANTRELL, PAUL R.
; REGISTRATION NUMBER: 36,470
; REFERENCE/DOCKET NUMBER: P9113
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3885
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4550 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-804-198-2

Query Match 3.7%; Score 8; DB 2; Length 4550;

Best Local Similarity 100.0%; Pred. No. 59;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 AEPAAAP 138
Db 2625 AEPAAAP 2632
|||||

RESULT 9

US-09-248-588-27
; Sequence 27, Application US/09248588
; Patent No. 6231864
; GENERAL INFORMATION:
; APPLICANT: Birkett, Ashley J.
; TITLE OF INVENTION: Strategically Modified Hepatitis B Core Proteins and
; TITLE OF INVENTION: their Derivatives
; FILE REFERENCE: SYN-101 4564/60529
; CURRENT APPLICATION NUMBER: US/09/248,588
; CURRENT FILING DATE: 1999-02-11
; EARLIER APPLICATION NUMBER: 60/074537
; EARLIER FILING DATE: 1998-02-12
; NUMBER OF SEQ ID NOS: 113
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 27
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Cryptosporidium parvum
; PUBLICATION INFORMATION:
; PATENT DOCUMENT NUMBER: WO 98/07320
US-09-248-588-27

Query Match 3.2%; Score 7; DB 4; Length 31;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 AAEPAAE 161
Db 14 AAEPAAE 20
|||||

RESULT 10

5273901-11
; Patent No. 5273901
; APPLICANT: JACOBSON, JAMES W.; STRAUSBERG, ROBERT L.; WILSON,
; SUSAN D.; POPE, SHARON H.; STRAUSBERG, SUSAN L.; RUFF, MICHAEL D.;
; AUGUSTINE, PATRICIA C.; DANFORTH, HARRY D.
; TITLE OF INVENTION: GENETICALLY ENGINEERED COCCIDIOSIS
; SPOOROZOITE 21.5 KB ANTIGEN, AC-6B
; NUMBER OF SEQUENCES: 11
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/581,693
; FILING DATE: 12-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 215,162
; FILING DATE: 05-JUL-1988
; APPLICATION NUMBER: 746,520
; FILING DATE: 19-JUN-1985
; APPLICATION NUMBER: 627,811
; FILING DATE: 05-JUL-1984
; SEQ ID NO:11:
; LENGTH:76
5273901-11

Query Match 3.2%; Score 7; DB 6; Length 76;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 AAAAPAA 126
Db 20 AAAAPAA 26
|||||

RESULT 11

5482709-10
; Patent No. 5482709
; APPLICANT: JACOBSON, JAMES W.; STRAUSBERG, ROBERT L.; WILSON,
; SUSAN D.; POPE, SHARON H.; STRAUSBERG, SUSAN L.; RUFF, MICHAEL D.;
; AUGUSTINE, PATRICIA C.; DANFORTH, HARRY D.
; TITLE OF INVENTION: EIMERIA ANTIGENIC COMPOSITION WHICH
; ELICITS ANTIBODIES AGAINST AVIAN COCCIDIOSIS
; NUMBER OF SEQUENCES: 10
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/148,432
; FILING DATE: 08-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 581,693
; FILING DATE: 12-SEP-1990
; APPLICATION NUMBER: 215,162
; FILING DATE: 05-JUL-1989
; APPLICATION NUMBER: 746,520
; FILING DATE: 19-JUN-1985
; APPLICATION NUMBER: 627,811
; FILING DATE: 05-JUL-1984
; SEQ ID NO:10:
; LENGTH: 76
5482709-10

Query Match 3.2%; Score 7; DB 6; Length 76;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 AAAAPAA 126
Db 20 AAAAPAA 26
|||||

RESULT 12

5496550-6
; Patent No. 5496550
; APPLICANT: WALLACH, MICHAEL; PUGATSCH, THEA; MENCHER, DAVID
; TITLE OF INVENTION: METHOD OF REDUCING THE OUTPUT OF EIMERIA
; OOCYSTS FROM A NEWBORN CHICK
; NUMBER OF SEQUENCES: 10
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/108,763
; FILING DATE: 17-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 642,219
; FILING DATE: 16-JAN-1991
; APPLICATION NUMBER: 310,603
; FILING DATE: 14-FEB-1989
; APPLICATION NUMBER: 155,245
; FILING DATE: 12-FEB-1988
; APPLICATION NUMBER: 896,611
; FILING DATE: 14-AUG-1986
; SEQ ID NO:6:
; LENGTH: 76
5496550-6

Query Match 3.2%; Score 7; DB 6; Length 76;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 AAAAPAA 126
Db 24 AAAAPAA 30
|||||

RESULT 13

US-08-301-162-10
; Sequence 10, Application US/08301162
; Patent No. 6022546
; GENERAL INFORMATION:
; APPLICANT: Knapp, Stefan
; APPLICANT: Ziegelmaier, Robert
; APPLICANT: Kupper, Hans

;
; TITLE OF INVENTION: Toxoplasma Gondii Antigens, The
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/301,162
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/167,128
; FILING DATE:
; APPLICATION NUMBER: US 07/623,086
; FILING DATE: 06-DEC-1990
; APPLICATION NUMBER: DE P3940598.2
; FILING DATE: 08-DEC-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Fleshner, Raz E.
; REGISTRATION NUMBER: 34,331
; REFERENCE/DOCKET NUMBER: 02481.1005-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4400
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 118 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-301-162-10

Query Match 3.2%; Score 7; DB 3; Length 118;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 120 AAAAPAA 126
Db 84 AAAAPAA 90
|||||

RESULT 14
US-09-461-240-10
; Sequence 10, Application US/09461240
; Patent No. 6326008
; GENERAL INFORMATION:
; APPLICANT: Knapp, Stefan
; Ziegelmaier, Robert
; Kupper, Hans
; TITLE OF INVENTION: Toxoplasma Gondii Antigens, The
; Preparation Thereof and the Use Thereof
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; Danner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/968,927
; FILING DATE: 03-Oct-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/301,162
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US/08/167,128
; FILING DATE: <Unknown>
; COMPUTER: IBM PC compatible

;
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/461,240
; FILING DATE: 16-Dec-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/301,162
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US/08/167,128
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 07/623,086
; FILING DATE: 06-DEC-1990
; APPLICATION NUMBER: DE P3940598.2
; FILING DATE: 08-DEC-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Fleshner, Raz E.
; REGISTRATION NUMBER: 34,331
; REFERENCE/DOCKET NUMBER: 02481.1005-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 118 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-461-240-10

Query Match 3.2%; Score 7; DB 4; Length 118;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 120 AAAAPAA 126
Db 84 AAAAPAA 90
|||||

RESULT 15
US-09-968-927-10
; Sequence 10, Application US/09968927
; Patent No. 6419925
; GENERAL INFORMATION:
; APPLICANT: Knapp, Stefan
; Ziegelmaier, Robert
; Kupper, Hans
; TITLE OF INVENTION: Toxoplasma Gondii Antigens, The
; Preparation Thereof and the Use Thereof
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; Danner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/968,927
; FILING DATE: 03-Oct-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/301,162
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US/08/167,128
; FILING DATE: <Unknown>
; COMPUTER: IBM PC compatible

;; APPLICATION NUMBER: US 07/623,086
;; FILING DATE: 06-DEC-1990
;; APPLICATION NUMBER: DE P3940598.2
;; FILING DATE: 08-DEC-1989
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Fleisher, Raz E.
;; REGISTRATION NUMBER: 34,331
;; REFERENCE/DOCKET NUMBER: 02481.1005-00000
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-408-4000
;; TELEFAX: 202-408-4400
;; INFORMATION FOR SEQ ID NO: 10:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 118 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-968-927-10

Query Match 3.2%; Score 7; DB 4; Length 118;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 AAAPAA 126
| | | | |
Db 84 AAAPAA 90

Search completed: July 7, 2003, 15:13:34
Job time : 36 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2003, 15:10:39 ; Search time 52 Seconds
(without alignments)
484.297 Million cell updates/sec

Title: US-09-923-236-2
Perfect score: 219
Sequence: 1 MKLLWACIVCVAFARKRRF.....TAKPAAPHPSPSLEQANQ 219

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 440863 seqs, 114992915 residues

Word size : 0
Total number of hits satisfying chosen parameters: 440863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PublishedApplications.AA.*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PTCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	219	100.0	219	9	US-09-922-469-2
2	219	100.0	219	10	US-09-922-480-2
3	219	100.0	219	10	US-09-923-236-2
4	140	63.9	219	9	US-09-746-783-106
5	9	4.1	798	9	US-10-156-761-13162
6	9	4.1	2969	9	US-09-738-626-4434
7	8	3.7	206	9	US-10-219-220-250
8	8	3.7	383	10	US-09-788-345-10
9	8	3.7	412	10	US-09-788-345-12
10	8	3.7	428	9	US-09-906-514-4
11	8	3.7	454	9	US-10-156-761-13939
12	8	3.7	511	10	US-09-864-761-34590
13	7	3.2	21	9	US-10-124-800-15
14	7	3.2	27	9	US-09-974-879-306
15	7	3.2	27	9	US-09-305-736-306
16	7	3.2	48	10	US-09-739-907-177
17	7	3.2	60	10	US-09-864-761-33554
18	7	3.2	60	10	US-09-864-761-33624
19	7	3.2	60	10	US-09-864-761-34127

20	7	3.2	61	10	US-09-864-761-39187	Sequence 39187, A
21	7	3.2	65	10	US-09-864-761-36194	Sequence 36194, A
22	7	3.2	78	9	US-09-820-843A-89	Sequence 89, Appl
23	7	3.2	88	10	US-09-764-853-503	Sequence 503, App
24	7	3.2	91	10	US-09-867-550-738	Sequence 738, App
25	7	3.2	95	10	US-09-739-907-178	Sequence 178, App
26	7	3.2	109	10	US-09-731-872-471	Sequence 471, App
27	7	3.2	136	10	US-09-764-853-720	Sequence 720, App
28	7	3.2	136	10	US-09-764-864-1234	Sequence 1234, Ap
29	7	3.2	145	10	US-09-764-853-476	Sequence 476, App
30	7	3.2	150	9	US-10-001-631-2	Sequence 2, Appli
31	7	3.2	156	10	US-09-815-242-12062	Sequence 12062, A
32	7	3.2	173	9	US-09-824-893A-258	Sequence 258, App
33	7	3.2	187	9	US-10-128-714-3496	Sequence 3496, Ap
34	7	3.2	187	9	US-10-128-714-8496	Sequence 8496, Ap
35	7	3.2	199	9	US-09-738-626-4417	Sequence 4417, Ap
36	7	3.2	199	10	US-09-953-260-2	Sequence 2, Appli
37	7	3.2	199	10	US-09-731-909A-2	Sequence 2, Appli
38	7	3.2	220	12	US-10-062-254-214	Sequence 214, App
39	7	3.2	241	9	US-10-156-761-11118	Sequence 11118, A
40	7	3.2	251	9	US-09-974-879-304	Sequence 304, App
41	7	3.2	251	9	US-09-305-736-304	Sequence 304, App
42	7	3.2	261	9	US-10-156-761-9177	Sequence 9177, Ap
43	7	3.2	285	9	US-10-156-761-13161	Sequence 13161, A
44	7	3.2	322	10	US-09-789-836-6	Sequence 6, Appli
45	7	3.2	345	9	US-10-060-036-183	Sequence 183, App

ALIGNMENTS

RESULT 1
US-09-922-469-2
; Sequence 2, Application US/09922469
; Patent No. US20020173027A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Adler, David A.
; TITLE OF INVENTION: SECRETED SALIVARY ZSIG63 POLYPEPTIDE
; FILE REFERENCE: 97-71
; CURRENT APPLICATION NUMBER: US/09/922,469
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/124,820
; PRIOR FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-922-469-2

Query Match	100.0%	Score 219;	DB 9;	Length 219;
Best Local Similarity	100.0%	Pred. No. 1.8e-191;		
Matches 219;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MKLLWACIVCVAFARKRRFPFGEDDNDGCHPLHPSLNIPYGINLPPPLYTRPVNTVP	60	
Db	1	MKLLWACIVCVAFARKRRFPFGEDDNDGCHPLHPSLNIPYGINLPPPLYTRPVNTVP	60	
Qy	61	SYPGNYTDTGLPSYFWILTSPGFYVYHIGRFFLATOLNYPPLPPRGFFPVPPSRFSA	120	
Db	61	SYPGNYTDTGLPSYFWILTSPGFYVYHIGRFFLATOLNYPPLPPRGFFPVPPSRFSA	120	
Qy	121	AAAPAAPPIAAEPAAAPLTATPVAAPAAAGPAAAPVAAEPAAAPVAAEPAAAPVAAEPAA	180	
Db	121	AAAPAAPPIAAEPAAAPLTATPVAAPAAAGPAAAPVAAEPAAAPVAAEPAAAPVAAEPAA	180	
Qy	181	EAPGVPEPAEPSPAPETAKPAAPHPSPSLEQANQ	219	
Db	181	EAPGVPEPAEPSPAPETAKPAAPHPSPSLEQANQ	219	

3

Db 121 AAAPAAPAAAPAAAPLT 140
|||||

RESULT 5

US-10-156-761-13162
; Sequence 13162, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 13162
; LENGTH: 798
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-13162

Query Match 4.1%: Score 9; DB 9; Length 798;
Best Local Similarity 100.0%; Pred. No. 6.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 198 PATAKPAAP 206
|||||
Db 29 PATAKPAAP 37

RESULT 6

US-09-738-626-4434
; Sequence 4434, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MEIKO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAORO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4434
; LENGTH: 2969
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4434

Query Match 4.1%: Score 9; DB 9; Length 2969;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 168 PAAEAPVAA 176
|||||
Db 1663 PAAEAPVAA 1671

RESULT 7

US-10-219-220-250
; Sequence 250, Application US/10219220
; Publication No. US20030082724A1
; GENERAL INFORMATION:
; APPLICANT: Flind, Barry
; APPLICANT: Lasham, Annette
; TITLE OF INVENTION: Compositions affecting programmed cell
; TITLE OF INVENTION: death and their use in the modification of plant development
; FILE REFERENCE: 11000.1022c1
; CURRENT APPLICATION NUMBER: US/10/219,220
; CURRENT FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: U.S. No. US20030082724A1 09/325,932
; PRIOR FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 250
; LENGTH: 206
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
US-10-219-220-250

Query Match 3.7%: Score 8; DB 9; Length 206;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 AAAAPAAP 127
|||||
Db 112 AAAAPAAP 119

RESULT 8

US-09-788-345-10
; Sequence 10, Application US/09788345
; Patent No. US20020147321A1
; GENERAL INFORMATION:
; APPLICANT: ALONSO BEDATE, Carlos
; APPLICANT: REQUENA ROLANIA, Jose M.
; APPLICANT: SOTO ALVAREZ, Manuel
; TITLE OF INVENTION: CHIMERIC GENE FORMED OF THE DNA SEQUENCES THAT ENCODE THE ANT1
; TITLE OF INVENTION: DETERMINANTS OF FOUR PROTEINS OF L. INFANTUM, USEFUL FOR SERC
; FILE REFERENCE: BEDATE-1A
; CURRENT APPLICATION NUMBER: US/09/788,345
; CURRENT FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 09/219,306
; PRIOR FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chimeric
US-09-788-345-10

Query Match 3.7%: Score 8; DB 10; Length 383;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 AAAAPAAP 127
|||||
Db 354 AAAAPAAP 361

```

RESULT 9
US-09-788-345-12
; Sequence 12, Application US/09788345
; Patent No. US20020147321A1
; GENERAL INFORMATION:
; APPLICANT: ALONSO BEDATE, Carlos
; APPLICANT: REQUENA ROLANIA, Jose M.
; APPLICANT: SOTO ALVAREZ, Manuel
; TITLE OF INVENTION: CHIMERIC GENE FORMED OF THE DNA SEQUENCES THAT ENCODE THE ANTIGEN
; TITLE OF INVENTION: DETERMINANTS OF FOUR PROTEINS OF L. INFANTUM, USEFUL FOR SEROLOG
; TITLE OF INVENTION: CANINE LEISHMANIOSIS AND PROTEIN OBTAINED
; FILE REFERENCE: BEDATE-1A
; CURRENT APPLICATION NUMBER: US/09/788,345
; CURRENT FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 09/219,306
; PRIOR FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: recombinant sequence from vectors pQ31 and pMal
US-09-788-345-12

Query Match 3.7%; Score 8; DB 10; Length 412;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 AAAAPAAP 127
DB 383 AAAAPAAP 390

RESULT 10
US-09-906-514-4
; Sequence 4, Application US/09906514
; Patent No. US20020170085A1
; GENERAL INFORMATION:
; APPLICANT: Kaeppler, Shawn
; APPLICANT: Springer, Nathan
; APPLICANT: Phillips, Ronald
; TITLE OF INVENTION: Methyl CpG Binding Domain Nucleic Acids from Maize
; FILE REFERENCE: Methybinding
; CURRENT APPLICATION NUMBER: US/09/906,514
; CURRENT FILING DATE: 2001-07-16
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Zea mays
US-09-906-514-4

Query Match 3.7%; Score 8; DB 9; Length 428;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 AEPAAAP 138
DB 292 AEPAAAP 299

RESULT 11
US-10-156-761-13939
; Sequence 13939, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO

```

```

; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 13939
; LENGTH: 454
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-13939

Query Match 3.7%; Score 8; DB 9; Length 454;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 150 AGAPVAAE 157
DB 332 AGAPVAAE 339

RESULT 12
US-09-864-761-34590
; Sequence 34590, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-x-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30

```


;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 34590
;; LENGTH: 511
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AB023049.1
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.7
;; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 9.3
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.9
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3
;; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 7.2
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 5.8
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.8
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.1
;; OTHER INFORMATION: SWISSPROT HIT: Q02817, EVALUE 3.00e-24
;; OTHER INFORMATION: EST_HUMAN HIT: AA487852.1, EVALUE 2.00e-73
US-09-864-761-34590

Query Match 3.7%; Score 8; DB 10; Length 511;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 204 AAPEPHPS 211
Db 95 AAPEPHPS 102
|||||

RESULT 13
US-10-124-800-15
;; Sequence 15, Application US/10124800
;; Publication No. US20020194641A1
;; GENERAL INFORMATION:
;; APPLICANT: Metz, James
;; APPLICANT: Barclay, William
;; APPLICANT: Platt, James
;; APPLICANT: Kuner, Jerry
;; TITLE OF INVENTION: PUFA Polyketide Synthase Systems and Uses Thereof
;; FILE REFERENCE: 2997-29
;; CURRENT APPLICATION NUMBER: US/10/124,800
;; CURRENT FILING DATE: 2002-04-16
;; PRIOR APPLICATION NUMBER: 09/231,899
;; PRIOR FILING DATE: 1999-01-14
;; PRIOR APPLICATION NUMBER: 60/284,066
;; PRIOR FILING DATE: 2001-04-16
;; PRIOR APPLICATION NUMBER: 60/298,796
;; PRIOR FILING DATE: 2001-06-15
;; PRIOR APPLICATION NUMBER: 60/323,269
;; PRIOR FILING DATE: 2001-09-18
;; NUMBER OF SEQ ID NOS: 37
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 15
;; LENGTH: 21
;; TYPE: PRT
;; ORGANISM: Schizochytrium sp.
US-10-124-800-15

Query Match 3.2%; Score 7; DB 9; Length 21;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 AAAPAAP 127
Db 7 AAAPAAP 13
|||||

RESULT 14
US-09-974-879-306
;; Sequence 306, Application US/09974879
;; Publication No. US20030028003A1
;; GENERAL INFORMATION:
;; APPLICANT: Rosen et al.
;; TITLE OF INVENTION: 125 Human Secreted Proteins
;; FILE REFERENCE: PZ020P2
;; CURRENT APPLICATION NUMBER: US/09/974,879
;; CURRENT FILING DATE: 2001-10-12
;; PRIOR APPLICATION NUMBER: US 60/239,893
;; PRIOR FILING DATE: 2000-10-13
;; PRIOR APPLICATION NUMBER: US 09/818,683
;; PRIOR FILING DATE: 2001-03-28
;; PRIOR APPLICATION NUMBER: US 09/305,736
;; PRIOR FILING DATE: 1999-05-05
;; PRIOR APPLICATION NUMBER: PCT/US98/23435
;; PRIOR FILING DATE: 1998-11-04
;; PRIOR APPLICATION NUMBER: US 60/064,911
;; PRIOR FILING DATE: 1997-11-07
;; PRIOR APPLICATION NUMBER: US 60/064,912
;; PRIOR FILING DATE: 1997-11-07
;; PRIOR APPLICATION NUMBER: US 60/064,983
;; PRIOR FILING DATE: 1997-11-07
;; PRIOR APPLICATION NUMBER: US 60/064,900
;; PRIOR FILING DATE: 1997-11-07
;; PRIOR APPLICATION NUMBER: US 60/064,988
;; PRIOR FILING DATE: 1997-11-07
;; PRIOR APPLICATION NUMBER: US 60/064,987
;; PRIOR FILING DATE: 1997-11-07
;; PRIOR APPLICATION NUMBER: US 60/064,908
;; PRIOR FILING DATE: 1997-11-07
;; PRIOR APPLICATION NUMBER: US 60/064,984
;; PRIOR FILING DATE: 1997-11-07
;; PRIOR APPLICATION NUMBER: US 60/064,985
;; PRIOR FILING DATE: 1997-11-07
;; PRIOR APPLICATION NUMBER: US 60/066,094
;; PRIOR FILING DATE: 1997-11-17
;; PRIOR APPLICATION NUMBER: US 60/066,100
;; PRIOR FILING DATE: 1997-11-17
;; PRIOR APPLICATION NUMBER: US 60/066,089
;; PRIOR FILING DATE: 1997-11-17
;; PRIOR APPLICATION NUMBER: US 60/066,095
;; PRIOR FILING DATE: 1997-11-17
;; PRIOR APPLICATION NUMBER: US 60/066,090
;; PRIOR FILING DATE: 1997-11-17
;; NUMBER OF SEQ ID NOS: 611
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 306
;; LENGTH: 27
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-974-879-306

Query Match 3.2%; Score 7; DB 9; Length 27;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 200 TAKPAAP 206
Db 1 TAKPAAP 7
|||||

RESULT 15
US-09-305-736-306
;; Sequence 306, Application US/09305736
;; Publication No. US2003008078A1
;; GENERAL INFORMATION:
;; APPLICANT: Feng et al.
;; TITLE OF INVENTION: 125 Human Secreted Proteins
;; FILE REFERENCE: PZ020P1

; CURRENT APPLICATION NUMBER: US/09/305,736
; CURRENT FILING DATE: 1999-05-05
; EARLIER APPLICATION NUMBER: PCT/US98/23435
; EARLIER FILING DATE: 1998-11-04
; EARLIER APPLICATION NUMBER: 60/064,911
; EARLIER FILING DATE: 1997-11-07
; EARLIER APPLICATION NUMBER: 60/064,912
; EARLIER FILING DATE: 1997-11-07
; EARLIER APPLICATION NUMBER: 60/064,983
; EARLIER FILING DATE: 1997-11-07
; EARLIER APPLICATION NUMBER: 60/064,900
; EARLIER FILING DATE: 1997-11-07
; EARLIER APPLICATION NUMBER: 60/064,988
; EARLIER FILING DATE: 1997-11-07
; EARLIER APPLICATION NUMBER: 60/064,987
; EARLIER FILING DATE: 1997-11-07
; EARLIER APPLICATION NUMBER: 60/064,908
; EARLIER FILING DATE: 1997-11-07
; EARLIER APPLICATION NUMBER: 60/064,984
; EARLIER FILING DATE: 1997-11-07
; EARLIER APPLICATION NUMBER: 60/064,985
; EARLIER FILING DATE: 1997-11-07
; EARLIER APPLICATION NUMBER: 60/066,094
; EARLIER FILING DATE: 1997-11-17
; EARLIER APPLICATION NUMBER: 60/066,100
; EARLIER FILING DATE: 1997-11-17
; EARLIER APPLICATION NUMBER: 60/066,089
; EARLIER FILING DATE: 1997-11-17
; EARLIER APPLICATION NUMBER: 60,066,095
; EARLIER FILING DATE: 1997-11-17
; EARLIER APPLICATION NUMBER: 60/066,090
; EARLIER FILING DATE: 1997-11-17
; NUMBER OF SEQ ID NOS: 612
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 306
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-305-736-306

Query Match 3.2%; Score 7; DB 9; Length 27;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 200 TAKPAAP 206
Db 1 TAKPAAP 7
|||||||

Search completed: July 7, 2003, 15:19:43
Job time : 53 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2003, 15:09:19 ; Search time 309 Seconds
(without alignments)
456.947 Million cell updates/sec

Title: US-09-923-236-2
Perfect score: 219
Sequence: 1 MKLLWACIVCAFAKRRF.....TAKPAPEHPSPSLQANQ 219

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 4569144 seqs, 644733110 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Pending_Patents_AA_Main:*

1: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US09_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US09_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US09_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US09_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US09_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US09_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US09_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US09_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US09_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US09_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US09_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	219	100.0	219	1	PCT-US02-08123-1081
2	219	100.0	219	1	PCT-US02-08277-723
3	219	100.0	219	1	PCT-US02-08278-1040
4	219	100.0	219	23	US-09-923-469-2
5	219	100.0	219	23	US-09-923-480-2
6	219	100.0	219	23	US-09-923-236-2

7	219	100.0	219	24	US-10-009-328-29	Sequence 29, Appl
8	219	100.0	219	27	US-60-162-287-4	Sequence 4, Appl
9	150	68.5	220	1	PCT-US02-08123-1632	Sequence 1632, Ap
10	150	68.5	220	1	PCT-US02-08277-1086	Sequence 1086, Ap
11	150	68.5	220	1	PCT-US02-08278-1563	Sequence 1563, Ap
12	150	68.5	221	1	PCT-US98-22376-190	Sequence 190, App
13	150	68.5	221	16	US-09-296-622-192	Sequence 192, App
14	150	68.5	221	26	US-10-231-417-192	Sequence 192, App
15	140	63.9	219	1	PCT-US98-13813-10	Sequence 10, Appl
16	140	63.9	219	13	US-08-958-304-10	Sequence 10, Appl
17	140	63.9	219	21	US-09-748-783-106	Sequence 106, App
18	126	57.5	143	22	US-09-834-366-18049	Sequence 18049, A
19	126	57.5	143	27	US-60-197-873-18049	Sequence 18049, A
20	125	57.1	143	20	US-09-621-976-5226	Sequence 5226, Ap
21	125	57.1	143	27	US-60-147-499-5226	Sequence 5226, Ap
22	112	51.1	112	12	US-08-887-195-14	Sequence 14, Appl
23	109	49.8	120	20	US-09-621-976-3905	Sequence 3905, Ap
24	109	49.8	120	27	US-60-147-499-3905	Sequence 3905, Ap
25	58	26.5	91	22	US-09-834-366-13475	Sequence 13475, A
26	58	26.5	91	27	US-60-197-873-13475	Sequence 13475, A
27	10	4.6	111	21	US-09-708-427-72642	Sequence 72642, A
28	9	4.1	75	20	US-09-614-150-19398	Sequence 19398, A
29	9	4.1	75	27	US-60-167-217-10318	Sequence 10318, A
30	9	4.1	75	27	US-60-167-217-19534	Sequence 19534, A
31	9	4.1	75	27	US-60-173-464-15968	Sequence 15968, A
32	9	4.1	75	27	US-60-191-681-15371	Sequence 15371, A
33	9	4.1	334	16	US-09-270-767-41733	Sequence 41733, A
34	9	4.1	697	27	US-60-161-932-2484	Sequence 2484, Ap
35	9	4.1	728	20	US-09-614-150-23082	Sequence 23082, A
36	9	4.1	728	21	US-09-791-537-144373	Sequence 144373, A
37	9	4.1	728	27	US-60-173-464-19004	Sequence 19004, A
38	9	4.1	728	27	US-60-191-637-23165	Sequence 23165, A
39	9	4.1	728	27	US-60-191-681-18263	Sequence 18263, A
40	9	4.1	864	21	US-09-791-537-48380	Sequence 48380, A
41	9	4.1	1874	20	US-09-602-7874-46	Sequence 46, Appl
42	9	4.1	2969	21	US-09-738-626-4434	Sequence 4434, Ap
43	8	3.7	65	1	PCT-US99-22855-2421	Sequence 2421, Ap
44	8	3.7	65	19	US-09-513-996A-32499	Sequence 32499, A
45	8	3.7	65	19	US-09-513-996A-32499	Sequence 32499, A

ALIGNMENTS

RESULT 1
PCT-US02-08123-1081
; Sequence 1081, Application PC/TUS0208123
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS904PCT
; CURRENT APPLICATION NUMBER: PCT/US02/08123
; CURRENT FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: US 60/331,287
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/306,171
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: US 60/277,340
; PRIOR FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 2048
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1081
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-08123-1081

Query Match 100.0% Score 219; DB 1; Length 219;
Best Local Similarity 100.0%; Pred. No. 6e-193;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKLLWACIVCAFAKRRFFIGEDNDNDGHPHPSLNIPYGRNLPPLPYRPNVTP 60
|||||


```
; Sequence 2, Application US/09922480
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Adler, David A.
; TITLE OF INVENTION: SECRETED SALIVARY ZSIG63 POLYPEPTIDE
; FILE REFERENCE: 97-71
; CURRENT APPLICATION NUMBER: US/09/922,480
; PRIOR APPLICATION NUMBER: 2001-08-03
; PRIOR FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-922-480-2

Query Match      100.0%; Score 219; DB 23; Length 219;
Best Local Similarity 100.0%; Pred. No. 6e-193;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKLLWACIVCVAFARKRRPFIFGEDDDGHPHLSLNIPYGINLPPPLYRPNVTVP 60
Db 1 MKLLWACIVCVAFARKRRPFIFGEDDDGHPHLSLNIPYGINLPPPLYRPNVTVP 60
QY 61 SYPGNTYTDGLSPYWLTSFGPPVYVYHIGRFPPLATQNLNVPPLPRGFFVPPSRFFSA 120
Db 61 SYPGNTYTDGLSPYWLTSFGPPVYVYHIGRFPPLATQNLNVPPLPRGFFVPPSRFFSA 120
QY 121 AAAPAAPPIAAEPAAPAAAPLTATPVAAEPAAGAPVAAEPAEAPVGAEPAAEAPVAAEPAA 180
Db 121 AAAPAAPPIAAEPAAPAAAPLTATPVAAEPAAGAPVAAEPAEAPVGAEPAAEAPVAAEPAA 180
QY 181 EAPVGEPAEAEPSAPAEPAATAKPAAPAEHPHSPSLEQANQ 219
Db 181 EAPVGEPAEAEPSAPAEPAATAKPAAPAEHPHSPSLEQANQ 219

RESULT 6
US-09-923-236-2
; Sequence 2, Application US/099223236
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Adler, David A.
; TITLE OF INVENTION: SECRETED SALIVARY ZSIG63 POLYPEPTIDE
; FILE REFERENCE: 97-71
; CURRENT APPLICATION NUMBER: US/09/923,236
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/124,820
; PRIOR FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-923-236-2

Query Match      100.0%; Score 219; DB 23; Length 219;
Best Local Similarity 100.0%; Pred. No. 6e-193;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKLLWACIVCVAFARKRRPFIFGEDDDGHPHLSLNIPYGINLPPPLYRPNVTVP 60
Db 1 MKLLWACIVCVAFARKRRPFIFGEDDDGHPHLSLNIPYGINLPPPLYRPNVTVP 60
QY 61 SYPGNTYTDGLSPYWLTSFGPPVYVYHIGRFPPLATQNLNVPPLPRGFFVPPSRFFSA 120
Db 61 SYPGNTYTDGLSPYWLTSFGPPVYVYHIGRFPPLATQNLNVPPLPRGFFVPPSRFFSA 120
QY 121 AAAPAAPPIAAEPAAPAAAPLTATPVAAEPAAGAPVAAEPAEAPVGAEPAAEAPVAAEPAA 180
Db 121 AAAPAAPPIAAEPAAPAAAPLTATPVAAEPAAGAPVAAEPAEAPVGAEPAAEAPVAAEPAA 180
QY 181 EAPVGEPAEAEPSAPAEPAATAKPAAPAEHPHSPSLEQANQ 219
Db 181 EAPVGEPAEAEPSAPAEPAATAKPAAPAEHPHSPSLEQANQ 219

RESULT 7
US-10-009-328-29
; Sequence 29, Application US/10009328
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: LAL, Preeti
; APPLICANT: YANG, Junming
; APPLICANT: YUE, Henry
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: TANG, Y. Tom
; APPLICANT: BANDMAN, Olga
; APPLICANT: BURFORD, Neil
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: PATTERSON, Chandra
; TITLE OF INVENTION: HUMAN TRANSPORT PROTEINS
; FILE REFERENCE: PF-0709 PCT
; CURRENT APPLICATION NUMBER: US/10/009,328
; CURRENT FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/139,923; 60/148,177; 60/149,357; 60/162,287
; PRIOR FILING DATE: 1999-06-17; 1999-08-10; 1999-08-18; 1999-10-28
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: PERL Program
; SEQ ID NO 29
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 1325518CD1
US-10-009-328-29

Query Match      100.0%; Score 219; DB 24; Length 219;
Best Local Similarity 100.0%; Pred. No. 6e-193;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKLLWACIVCVAFARKRRPFIFGEDDDGHPHLSLNIPYGINLPPPLYRPNVTVP 60
Db 1 MKLLWACIVCVAFARKRRPFIFGEDDDGHPHLSLNIPYGINLPPPLYRPNVTVP 60
QY 61 SYPGNTYTDGLSPYWLTSFGPPVYVYHIGRFPPLATQNLNVPPLPRGFFVPPSRFFSA 120
Db 61 SYPGNTYTDGLSPYWLTSFGPPVYVYHIGRFPPLATQNLNVPPLPRGFFVPPSRFFSA 120
QY 121 AAAPAAPPIAAEPAAPAAAPLTATPVAAEPAAGAPVAAEPAEAPVGAEPAAEAPVAAEPAA 180
Db 121 AAAPAAPPIAAEPAAPAAAPLTATPVAAEPAAGAPVAAEPAEAPVGAEPAAEAPVAAEPAA 180
QY 181 EAPVGEPAEAEPSAPAEPAATAKPAAPAEHPHSPSLEQANQ 219
Db 181 EAPVGEPAEAEPSAPAEPAATAKPAAPAEHPHSPSLEQANQ 219

RESULT 8
US-60-162-287-4
; Sequence 4, Application US/60162287
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Tang, Y. Tom
; APPLICANT: Burford, Neil
; APPLICANT: Yue, Henry
; APPLICANT: Lal, Preeti
; APPLICANT: Patterson, Chandra
; APPLICANT: Baughn, Mariah R.
```

```
; APPLICANT: Lu, Dzung Aina M.
; TITLE OF INVENTION: TRANSPORT PROTEINS
; FILE REFERENCE: PF-0748 P
; CURRENT APPLICATION NUMBER: US/60/162,287
; CURRENT FILING DATE: 1999-10-28
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PERL Program
; SEQ ID NO 4
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 1325518CD1
US-60-162-287-4

Query Match      100.0%; Score 219; DB 27; Length 219;
Best Local Similarity 100.0%; Pred. No. 6e-193;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKLLWACIVCVAFARKRRPFIFGDDNDGHPHSLNIPYGINLPPPLYRPVNTVP 60
Db 1 MKLLWACIVCVAFARKRRPFIFGDDNDGHPHSLNIPYGINLPPPLYRPVNTVP 60
Qy 61 SYPGNTYDTGLSPYFWILTSPPGYVYHIRGFPLATQLNVPLPPRGFFVPPSRFFSA 120
Db 61 SYPGNTYDTGLSPYFWILTSPPGYVYHIRGFPLATQLNVPLPPRGFFVPPSRFFSA 120
Qy 121 AAAPAAPPIAAEPAAPLTAATPVAAEPAAEPVGAEPAAEPVAAEPAA 180
Db 121 AAAPAAPPIAAEPAAPLTAATPVAAEPAAEPVGAEPAAEPVAAEPAA 180
Qy 181 EAPGVPEAAEPSPAEPAEPATAKPAEPHPSPSLEQANQ 219
Db 181 EAPGVPEAAEPSPAEPAEPATAKPAEPHPSPSLEQANQ 219

RESULT 9
PCT-US02-08123-1632
; Sequence 1632, Application PC/TUS0208123
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS904PCT
; CURRENT APPLICATION NUMBER: PCT/US02/08123
; CURRENT FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: US 60/331,287
; PRIOR FILING DATE: 2001-11-13
; PRIOR FILING DATE: 2001-07-19
; PRIOR FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 2048
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1632
; LENGTH: 220
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (159)
; OTHER INFORMATION: Xaa equals any amino acid
PCT-US02-08123-1632

Query Match      68.5%; Score 150; DB 1; Length 220;
Best Local Similarity 100.0%; Pred. No. 2.3e-129;
Matches 150; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKLLWACIVCVAFARKRRPFIFGDDNDGHPHSLNIPYGINLPPPLYRPVNTVP 60
Db 1 MKLLWACIVCVAFARKRRPFIFGDDNDGHPHSLNIPYGINLPPPLYRPVNTVP 60
Qy 61 SYPGNTYDTGLSPYFWILTSPPGYVYHIRGFPLATQLNVPLPPRGFFVPPSRFFSA 120
Db 61 SYPGNTYDTGLSPYFWILTSPPGYVYHIRGFPLATQLNVPLPPRGFFVPPSRFFSA 120
Qy 121 AAAPAAPPIAAEPAAPLTAATPVAAEPAA 150
Db 121 AAAPAAPPIAAEPAAPLTAATPVAAEPAA 150

RESULT 11
PCT-US02-08278-1563
; Sequence 1563, Application PC/TUS0208278
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS902PCT
; CURRENT APPLICATION NUMBER: PCT/US02/08278
; CURRENT FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: US 60/331,287
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/306,171
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: US 60/277,340
; NUMBER OF SEQ ID NOS: 1988
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1563
; LENGTH: 220
; TYPE: PRT
; ORGANISM: Homo sapiens
```

```
Db 61 SYPGNTYDTGLSPYFWILTSPPGYVYHIRGFPLATQLNVPLPPRGFFVPPSRFFSA 120
Qy 121 AAAPAAPPIAAEPAAPLTAATPVAAEPAA 150
Db 121 AAAPAAPPIAAEPAAPLTAATPVAAEPAA 150

RESULT 10
PCT-US02-08277-1086
; Sequence 1086, Application PC/TUS0208277
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS907PCT
; CURRENT APPLICATION NUMBER: PCT/US02/08277
; CURRENT FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: US 60/331,287
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/306,171
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: US 60/277,340
; NUMBER OF SEQ ID NOS: 1357
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1086
; LENGTH: 220
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (159)
; OTHER INFORMATION: Xaa equals any amino acid
PCT-US02-08277-1086

Query Match      68.5%; Score 150; DB 1; Length 220;
Best Local Similarity 100.0%; Pred. No. 2.3e-129;
Matches 150; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKLLWACIVCVAFARKRRPFIFGDDNDGHPHSLNIPYGINLPPPLYRPVNTVP 60
Db 1 MKLLWACIVCVAFARKRRPFIFGDDNDGHPHSLNIPYGINLPPPLYRPVNTVP 60
Qy 61 SYPGNTYDTGLSPYFWILTSPPGYVYHIRGFPLATQLNVPLPPRGFFVPPSRFFSA 120
Db 61 SYPGNTYDTGLSPYFWILTSPPGYVYHIRGFPLATQLNVPLPPRGFFVPPSRFFSA 120
Qy 121 AAAPAAPPIAAEPAAPLTAATPVAAEPAA 150
Db 121 AAAPAAPPIAAEPAAPLTAATPVAAEPAA 150

RESULT 11
PCT-US02-08278-1563
; Sequence 1563, Application PC/TUS0208278
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS902PCT
; CURRENT APPLICATION NUMBER: PCT/US02/08278
; CURRENT FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: US 60/331,287
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/306,171
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: US 60/277,340
; NUMBER OF SEQ ID NOS: 1988
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1563
; LENGTH: 220
; TYPE: PRT
; ORGANISM: Homo sapiens
```

```

; FEATURE:
; NAME/KEY: SITE
; LOCATION: (159)
; OTHER INFORMATION: xaa equals any amino acid
PCT-US02-08278-1563

```

Query Match 68.5%; Score 150; DB 1; Length 220;

QY	1	MKLLWACIVCVAFARRRFF	IGEDDNDGDHPLHPSLNIPY	GIRNLPPLYYRPVNTVP	60
Db	1	MKLLWACIVCVAFARRRFF	IGEDDNDGDHPLHPSLNIPY	GIRNLPPLYYRPVNTVP	60
QY	61	SYPGNTYDTGLSPY	WILTS	SPGPPYYH	IRGFPLATQLNV
Db	61	SYPGNTYDTGLSPY	WILTS	SPGPPYYH	IRGFPLATQLNV
QY	121	AAAPPAAPPIAAE	PAAPAA	PLATPVAA	EPAA 150
Db	121	AAAPPAAPPIAAE	PAAPAA	PLATPVAA	EPAA 150

```

RESULT 12
PC/TUS9822376-190
; Sequence 190, Application PC/TUS9822376
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: 148 Human Secreted Proteins
; FILE REFERENCE: P2019.PCT
; CURRENT APPLICATION NUMBER: PCT/US98/22376
; CURRENT FILING DATE: 1998-10-23
; EARLIER APPLICATION NUMBER: 60/063,099
; EARLIER FILING DATE: 1997-10-24
; EARLIER APPLICATION NUMBER: 60/063,088
; EARLIER FILING DATE: 1997-10-24
; EARLIER APPLICATION NUMBER: 60/063,387
; EARLIER FILING DATE: 1997-10-24
; EARLIER APPLICATION NUMBER: 60/063,148
; EARLIER FILING DATE: 1997-10-24
; EARLIER APPLICATION NUMBER: 60/063,386
; EARLIER FILING DATE: 1997-10-24
; EARLIER APPLICATION NUMBER: 60/062,784
; EARLIER FILING DATE: 1997-10-24
; EARLIER APPLICATION NUMBER: 60/063,091
; EARLIER FILING DATE: 1997-10-24
; EARLIER APPLICATION NUMBER: 60/063,090
; EARLIER FILING DATE: 1997-10-24
; EARLIER APPLICATION NUMBER: 60/063,089
; EARLIER FILING DATE: 1997-10-24
; EARLIER APPLICATION NUMBER: 60/063,092
; EARLIER FILING DATE: 1997-10-24
; EARLIER APPLICATION NUMBER: 60/063,111
; EARLIER FILING DATE: 1997-10-24
; EARLIER APPLICATION NUMBER: 60/063,101
; EARLIER FILING DATE: 1997-10-24
; EARLIER APPLICATION NUMBER: 60/063,109
; EARLIER FILING DATE: 1997-10-24
; EARLIER APPLICATION NUMBER: 60/063,110
; EARLIER FILING DATE: 1997-10-24
; EARLIER APPLICATION NUMBER: 60/063,097
; EARLIER FILING DATE: 1997-10-24
; EARLIER APPLICATION NUMBER: 60/063,100
; EARLIER FILING DATE: 1997-10-24
; EARLIER APPLICATION NUMBER: 60/063,098
; EARLIER FILING DATE: 1997-10-24
; NUMBER OF SEQ ID NOS: 492
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 190
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:

```

```

: NAME/KEY: SITE
: LOCATION: (159)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (221)
: OTHER INFORMATION: Xaa equals stop translation
PCT-US98-22376-190

```

[illegible]

```

RESULT 13
US-09-296-622-192
; Sequence 192, Application US/09296622
; GENERAL INFORMATION:
; APPLICANT: Feng et al.
; TITLE OF INVENTION: 148 Human Secreted Proteins
; CURRENT APPLICATION NUMBER: US/09/296,622
; EARLIEST FILING DATE: 1999-04-23
; EARLIER APPLICATION NUMBER: PCT/US98/22376
; EARLIER FILING DATE: 1998-10-23
; EARLIER APPLICATION NUMBER: 60/063,099
; EARLIER FILING DATE: 1997-10-24
; EARLIER APPLICATION NUMBER: 60/063,088
; EARLIER FILING DATE: 1997-10-24
; EARLIER APPLICATION NUMBER: 60/063,100
; EARLIER FILING DATE: 1997-10-24
; EARLIER APPLICATION NUMBER: 60/063,387
; EARLIER FILING DATE: 1997-10-24
; EARLIER APPLICATION NUMBER: 60/063,148
; EARLIER FILING DATE: 1997-10-24
; EARLIER APPLICATION NUMBER: 60/063,386
; EARLIER FILING DATE: 1997-10-24
; EARLIER APPLICATION NUMBER: 60/062,784
; EARLIER FILING DATE: 1997-10-24
; EARLIER APPLICATION NUMBER: 60/063,091
; EARLIER FILING DATE: 1997-10-24
; EARLIER APPLICATION NUMBER: 60/063,090
; EARLIER FILING DATE: 1997-10-24
; EARLIER APPLICATION NUMBER: 60/063,089
; EARLIER FILING DATE: 1997-10-24
; EARLIER APPLICATION NUMBER: 60/063,092
; EARLIER FILING DATE: 1997-10-24
; EARLIER APPLICATION NUMBER: 60/063,111
; EARLIER FILING DATE: 1997-10-24
; EARLIER APPLICATION NUMBER: 60/063,101
; EARLIER FILING DATE: 1997-10-24
; EARLIER APPLICATION NUMBER: 60/063,109
; EARLIER FILING DATE: 1997-10-24
; EARLIER APPLICATION NUMBER: 60/063,110
; EARLIER FILING DATE: 1997-10-24
; EARLIER APPLICATION NUMBER: 60/063,098
; EARLIER FILING DATE: 1997-10-24
; EARLIER APPLICATION NUMBER: 60/063,097
; EARLIER FILING DATE: 1997-10-24
; NUMBER OF SEQ ID NOS: 619
SOFTWARE: PatentV. 2.0

```

; SEQ ID NO 192
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: SITE
; LOCATION: (159)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (221)
; OTHER INFORMATION: Xaa equals stop translation
US-09-236-622-192

Query Match 68.5%; Score 150; DB 16; Length 221;
Best Local Similarity 100.0%; Pred. No. 2.3e-129; Mismatches 0; Indels 0; Gaps 0;
Matches 150; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKLLWACIVCVAFARKRRFFIGEDDNDGHLHPSLNIPYGINLPPLPPYRPVNTVP 60
DB 1 MKLLWACIVCVAFARKRRFFIGEDDNDGHLHPSLNIPYGINLPPLPPYRPVNTVP 60
QY 61 SYPGNTYDTGLSPYFWILTSFGPPYVYHIRGFLATQNLNVPPLPRGFFVPPSRFFSA 120
DB 61 SYPGNTYDTGLSPYFWILTSFGPPYVYHIRGFLATQNLNVPPLPRGFFVPPSRFFSA 120
QY 121 AAAPAAPPIAAEPAAPLTPVAEPAA 150
DB 121 AAAPAAPPIAAEPAAPLTPVAEPAA 150

RESULT 14
US-10-231-417-192
; Sequence 192, Application US/10231417
; GENERAL INFORMATION:
; APPLICANT: Feng et al.
; TITLE OF INVENTION: 148 Human Secreted Proteins
; FILE REFERENCE: P2019P1
; CURRENT APPLICATION NUMBER: US/10/231,417
; CURRENT FILING DATE: 2002-08-30
; PRIOR APPLICATION NUMBER: US/09/296,622
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 619
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 192
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: SITE
; LOCATION: (159)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (221)
; OTHER INFORMATION: Xaa equals stop translation
US-10-231-417-192

Query Match 68.5%; Score 150; DB 26; Length 221;
Best Local Similarity 100.0%; Pred. No. 2.3e-129; Mismatches 0; Indels 0; Gaps 0;
Matches 150; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKLLWACIVCVAFARKRRFFIGEDDNDGHLHPSLNIPYGINLPPLPPYRPVNTVP 60
DB 1 MKLLWACIVCVAFARKRRFFIGEDDNDGHLHPSLNIPYGINLPPLPPYRPVNTVP 60
QY 61 SYPGNTYDTGLSPYFWILTSFGPPYVYHIRGFLATQNLNVPPLPRGFFVPPSRFFSA 120
DB 61 SYPGNTYDTGLSPYFWILTSFGPPYVYHIRGFLATQNLNVPPLPRGFFVPPSRFFSA 120
QY 121 AAAPAAPPIAAEPAAPLTPVAEPAA 150
DB 121 AAAPAAPPIAAEPAAPLTPVAEPAA 150

RESULT 15
PCT-US98-13813-10
; Sequence 10, Application PC/TUS9813813
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavallie, Edward R.
; APPLICANT: Racie, Lisa A.
; APPLICANT: Treacy, Maurice
; APPLICANT: Spaulding, Vikki
; APPLICANT: Agostino, Michael J.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US98/13813
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sprunger, Suzanne A.
; REGISTRATION NUMBER: 41,323
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8284
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 219 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US98-13813-10

Query Match 63.9%; Score 140; DB 1; Length 219;
Best Local Similarity 100.0%; Pred. No. 3.7e-120; Mismatches 0; Indels 0; Gaps 0;
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKLLWACIVCVAFARKRRFFIGEDDNDGHLHPSLNIPYGINLPPLPPYRPVNTVP 60
DB 1 MKLLWACIVCVAFARKRRFFIGEDDNDGHLHPSLNIPYGINLPPLPPYRPVNTVP 60
QY 61 SYPGNTYDTGLSPYFWILTSFGPPYVYHIRGFLATQNLNVPPLPRGFFVPPSRFFSA 120
DB 61 SYPGNTYDTGLSPYFWILTSFGPPYVYHIRGFLATQNLNVPPLPRGFFVPPSRFFSA 120
QY 121 AAAPAAPPIAAEPAAPLTPVAEPAA 140
DB 121 AAAPAAPPIAAEPAAPLTPVAEPAA 140

Search completed: July 7, 2003, 15:18:44
Job time : 311 secs

GenCore version 5.1.6

Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2003, 15:12:14 ; Search time 110 Seconds
(without alignments)

519.970 Million cell updates/sec

Title: US-09-923-236-2

Perfect score: 219

Sequence: 1 MKLLWACIVCVAFARKRF.....TAKPAEPHPSPSLEQANQ 219

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1209432 seqs, 261172177 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1209432

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Pending_Patents_AA_New.*

- 1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep.*
- 2: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep4.*
- 3: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep.*
- 4: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep.*
- 5: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep.*
- 6: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep4.*
- 7: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep.*
- 8: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep4.*
- 9: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep.*
- 10: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep4.*
- 11: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep.*
- 12: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep4.*
- 13: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep.*
- 14: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	4.1	201	12	US-10-437-963-155644
2	9	4.1	365	12	US-10-437-963-116009
3	9	4.1	653	12	US-10-437-963-170175
4	9	4.1	798	12	US-10-156-761-13162
5	9	4.1	883	12	US-10-437-963-116012
6	8	3.7	91	12	US-10-282-122A-47349
7	8	3.7	115	12	US-10-437-963-179665
8	8	3.7	121	12	US-10-218-140-1314
9	8	3.7	125	12	US-10-437-963-156203
10	8	3.7	129	12	US-10-282-122A-54876
11	8	3.7	132	12	US-10-289-762-96
12	8	3.7	142	12	US-10-437-963-127079
13	8	3.7	149	12	US-10-425-114-40736
14	8	3.7	149	12	US-10-425-114-40752
15	8	3.7	149	12	US-10-425-114-40754
16	8	3.7	149	12	US-10-425-114-42294
17	8	3.7	149	12	US-10-425-114-53951
18	8	3.7	165	12	US-10-446-203-9067
19	8	3.7	168	12	US-10-366-683-18199

```
20      3.7      168      12      US-10-419-128-18199      Sequence 18199, A
21      3.7      177      12      US-10-437-963-140542      Sequence 140542, A
22      3.7      180      12      US-10-424-599-266399      Sequence 266399, A
23      3.7      190      12      US-10-366-683-17963      Sequence 17963, A
24      3.7      190      12      US-10-419-128-17963      Sequence 17963, A
25      3.7      191      12      US-10-437-963-118288      Sequence 118288, A
26      3.7      194      12      US-10-437-963-158973      Sequence 158973, A
27      3.7      233      12      US-10-437-963-129524      Sequence 129524, A
28      3.7      259      12      US-10-437-963-119942      Sequence 119942, A
29      3.7      305      12      US-10-424-599-181850      Sequence 181850, A
30      3.7      307      12      US-10-366-683-21588      Sequence 21588, A
31      3.7      307      12      US-10-419-128-21588      Sequence 21588, A
32      3.7      309      12      US-10-425-114-41494      Sequence 41494, A
33      3.7      314      12      US-10-437-963-119944      Sequence 119944, A
34      3.7      326      12      US-10-424-599-257117      Sequence 257117, A
35      3.7      340      12      US-10-437-963-119545      Sequence 119545, A
36      3.7      342      12      US-10-369-493-12383      Sequence 12383, A
37      3.7      344      12      US-10-437-963-204508      Sequence 204508, A
38      3.7      354      12      US-10-282-122A-49617      Sequence 49617, A
39      3.7      356      12      US-10-437-963-190346      Sequence 190346, A
40      3.7      367      12      US-10-369-493-7317      Sequence 7317, Ap
41      3.7      383      12      US-10-337-312-10      Sequence 10, Appl
42      3.7      410      12      US-10-366-683-30606      Sequence 30606, A
43      3.7      410      12      US-10-419-128-30606      Sequence 30606, A
44      3.7      411      12      US-10-437-963-198329      Sequence 198329, A
45      3.7      412      12      US-10-337-312-12      Sequence 12, Appl
```

ALIGNMENTS

```
RESULT 1
US-10-437-963-155644
; Sequence 155644, Application US/10437963
; GENERAL INFORMATION:
```

```
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
```

```
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
```

```
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
```

```
; SEQ ID NO 155644
```

```
; LENGTH: 201
```

```
; TYPE: PRT
```

```
; ORGANISM: Oryza sativa
```

```
; FEATURE:
```

```
; OTHER INFORMATION: Clone ID: PAT_MRT4530_55389C.1.pep
```

```
US-10-437-963-155644
```

```
Query Match      4.1%      Score 9;      DB 12;      Length 201;
Best Local Similarity 100.0%;      Pred. No. 4.1;
Matches      9;      Conservative      0;      Mismatches      0;      Indels      0;      Gaps      0;
```

```
QY      120      AAAAPAAPPP      128
```

```
|||||||
```

```
Db      12      AAAAPAAPPP      20
```

```
RESULT 2
```

```
US-10-437-963-116009
; Sequence 116009, Application US/10437963
; GENERAL INFORMATION:
```

```
; APPLICANT: La Rosa, Thomas J.
```

```
; APPLICANT: Kovalic, David K.
```

```
; APPLICANT: Zhou, Yihua
```

```
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; PRIOR FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 116009
; LENGTH: 365
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_19551C.1.pep
US-10-437-963-116009

Query Match
Best Local Similarity 4.1%; Score 9; DB 12; Length 365;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 120 AAAAPAAP 128
Db 343 AAAAPAAP 351

RESULT 3
US-10-437-963-170175
; Sequence 170175, Application US/10437963
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; PRIOR FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 170175
; LENGTH: 653
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_68524C.1.pep
US-10-437-963-170175

Query Match
Best Local Similarity 4.1%; Score 9; DB 12; Length 653;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 120 AAAAPAAP 128
Db 551 AAAAPAAP 559

RESULT 4
US-10-156-761-13162
; Sequence 13162, Application US/10156761
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
```

```
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; CURRENT APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 13162
; LENGTH: 798
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-13162

Query Match
Best Local Similarity 4.1%; Score 9; DB 12; Length 798;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 198 PATAKPAAP 206
Db 29 PATAKPAAP 37

RESULT 5
US-10-437-963-116012
; Sequence 116012, Application US/10437963
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 116012
; LENGTH: 883
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_19554C.1.pep
US-10-437-963-116012

Query Match
Best Local Similarity 4.1%; Score 9; DB 12; Length 883;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 120 AAAAPAAP 128
Db 755 AAAAPAAP 763

RESULT 6
US-10-282-122A-47349
; Sequence 47349, Application US/10282122A
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haseibeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
```

APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 47349
LENGTH: 91
TYPE: PRT
ORGANISM: Burkholderia cepacia
US-10-282-122A-47349

Query Match 3.7%; Score 8; DB 12; Length 91;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 AAAAPAP 127
DB 6 AAAAPAP 13

RESULT 7
US-10-437-963-179665
Sequence 179665, Application US/10437963
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 179665
LENGTH: 115
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_77104C.1.pap
US-10-437-963-179665

Query Match 3.7%; Score 8; DB 12; Length 115;
Best Local Similarity 100.0%; Pred. No. 21;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 119 SAAAAPAA 126
DB 86 SAAAAPAA 93

RESULT 8
US-10-218-140-1314
Sequence 1314, Application US/10218140
GENERAL INFORMATION:
APPLICANT: Leach, Martin D.
APPLICANT: Shimkets, Richard A.
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND POLYPEPTIDES
FILE REFERENCE: ENCODED THEREBY
FILE REFERENCE: 15966-543 CON
CURRENT APPLICATION NUMBER: US/10/218,140
CURRENT FILING DATE: 2002-08-12
PRIOR APPLICATION NUMBER: 09/540,763
PRIOR FILING DATE: 2000-03-30
PRIOR APPLICATION NUMBER: 60/127,728
PRIOR FILING DATE: 1999-04-05
PRIOR APPLICATION NUMBER: 60/127,636
PRIOR FILING DATE: 1999-04-02
PRIOR APPLICATION NUMBER: 60/127,607
PRIOR FILING DATE: 1999-03-31
NUMBER OF SEQ ID NOS: 6322
SOFTWARE: Curator Version 1.0
SEQ ID NO 1314
LENGTH: 121
TYPE: PRT
ORGANISM: Homo sapiens
US-10-218-140-1314

Query Match 3.7%; Score 8; DB 12; Length 121;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 AAAPAAP 128
DB 103 AAAPAAP 110

RESULT 9
US-10-437-963-156203
Sequence 156203, Application US/10437963
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 156203
LENGTH: 125
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_55894C.1.pap
US-10-437-963-156203

Query Match 3.7%; Score 8; DB 12; Length 125;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 AAAPAAP 128
| | | | | | | |
Db 11 AAAPAAP 18

RESULT 10

US-10-282-122A-54876
; Sequence 54876, Application US/10282122A

; GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.

; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A

; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20

; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09

; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09

; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 78614

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 54876

; LENGTH: 129

; TYPE: PRT

; ORGANISM: Chlamydia pneumoniae

US-10-282-122A-54876

Query Match

Best Local Similarity 3.7%; Score 8; DB 12; Length 129;

Mismatches 0; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 171 EAPVAAEP 178

| | | | | | | |

Db 50 EAPVAAEP 57

RESULT 11

US-10-289-762-96

; Sequence 96, Application US/10289762

; GENERAL INFORMATION:

; APPLICANT: Griffiths, R.

; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments

; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention

; TITLE OF INVENTION: and treatment of infection

; FILE REFERENCE: 9710-003-999

; CURRENT APPLICATION NUMBER: US/10/289,762

; CURRENT FILING DATE: 2003-03-27
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 96
; LENGTH: 132
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-10-289-762-96

Query Match

Best Local Similarity 3.7%; Score 8; DB 12; Length 132;

Mismatches 0; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 171 EAPVAAEP 178

| | | | | | | |

Db 53 EAPVAAEP 60

RESULT 12

US-10-437-963-127079

; Sequence 127079, Application US/10437963

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 127079

; LENGTH: 142

; TYPE: PRT

; ORGANISM: Oryza sativa

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT4530_29565C.1.pap

US-10-437-963-127079

Query Match

Best Local Similarity 3.7%; Score 8; DB 12; Length 142;

Mismatches 0; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 119 SAAAPAA 126

| | | | | | | |

Db 90 SAAAPAA 97

RESULT 13

US-10-425-114-40736

; Sequence 40736, Application US/10425114

; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong

; APPLICANT: Zhou, Yihua

; APPLICANT: Kovalic, David K.

; APPLICANT: Screen, Steven E.

; APPLICANT: Tabaska, Jack E.

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53313)B

; CURRENT APPLICATION NUMBER: US/10/425,114

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128

; SEQ ID NO 40736

; LENGTH: 149

; TYPE: PRT

; ORGANISM: Arabidopsis thaliana

; FEATURE:

; OTHER INFORMATION: Clone ID: LIB25-036-B4_FLI.pep
US-10-425-114-40736

Query Match 3.7%; Score 8; DB 12; Length 149;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 AAEPAAAA 137
Db 23 AAEPAAAA 30

RESULT 14

US-10-425-114-40752
; Sequence 40752, Application US/10425114
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 40752
; LENGTH: 149
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB25-102-G8_FLI.pep
US-10-425-114-40752

Query Match 3.7%; Score 8; DB 12; Length 149;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 AAEPAAAA 137
Db 23 AAEPAAAA 30

RESULT 15

US-10-425-114-40754
; Sequence 40754, Application US/10425114
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 40754
; LENGTH: 149
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB25-104-C11_FLI.pep
US-10-425-114-40754

Query Match 3.7%; Score 8; DB 12; Length 149;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 AAEPAAAA 137
Db 23 AAEPAAAA 30

Search completed: July 7, 2003, 15:21:41
Job time : 111 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2003, 15:07:49 ; Search time 40 Seconds
(without alignments)
526.336 Million cell updates/sec

Title: US-09-923-236-2

Perfect score: 219

Sequence: 1 MKLLWACIVCVAFARKRRF.....TAKPAPEPHSPSLEQANQ 219

Scoring table:

OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283224 seqs, 96134422 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

PIR_73:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	4.1	581	2 T36267	probable glutamyl-
2	9	4.1	864	1 JC1422	nitrate reductase
3	8	3.7	110	1 R6BYP3	60s acidic ribosom
4	8	3.7	129	2 H86500	L7/L12 ribosomal p
5	8	3.7	129	2 C72122	ribosomal protein
6	8	3.7	193	2 H86404	probable lipid tra
7	8	3.7	228	2 E87612	cytochrome c, memb
8	8	3.7	268	2 G70645	hypothetical prote
9	8	3.7	276	2 G95887	probable ABC trans
10	8	3.7	291	2 T2434	DNA binding protei
11	8	3.7	325	2 T48873	electron transfer
12	8	3.7	336	2 A87300	TPR domain protein
13	8	3.7	353	2 S41958	pupal cuticle pro
14	8	3.7	381	2 T46827	phenoxymethyl di
15	8	3.7	396	1 JH0633	cellular tumor ant
16	8	3.7	495	2 T52066	probable farnesyl
17	8	3.7	506	2 A40679	transcription enha
18	8	3.7	523	2 B40679	transcription enha
19	8	3.7	550	2 C75557	hypothetical prote
20	8	3.7	574	2 S23530	H+-transporting tw
21	8	3.7	611	2 C96030	probable cell divi
22	8	3.7	846	2 S2418	GTP-binding regula
23	8	3.7	1476	2 A45773	kelch protein, lon
24	8	3.7	4307	2 T20721	hypothetical prote
25	7	3.2	38	2 S23173	photosystem I chai
26	7	3.2	78	2 E83446	hypothetical prote
27	7	3.2	82	1 FDFLAW	antifreeze protein
28	7	3.2	82	2 S02326	antifreeze protein
29	7	3.2	82	2 A05161	antifreeze protein

30 7 3.2 82 2 I51125 antifreeze protein
31 7 3.2 100 2 F95852 hypothetical prote
32 7 3.2 101 2 H69091 ribosomal protein
33 7 3.2 102 2 D64363 ribosomal protein
34 7 3.2 106 1 R6DOP2 acidic ribosomal p
35 7 3.2 107 1 R6KMIC acidic ribosomal p
36 7 3.2 109 1 R6UTP1 acidic ribosomal p
37 7 3.2 110 2 T37490 ribosomal protein
38 7 3.2 111 2 E72524 probable ribosomal
39 7 3.2 111 2 E95857 hypothetical prote
40 7 3.2 112 1 R5FF2E acidic ribosomal p
41 7 3.2 112 2 S54179 acidic ribosomal p
42 7 3.2 112 2 E86141 protein T25K16.9
43 7 3.2 114 1 R5RT12 acidic ribosomal p
44 7 3.2 114 1 R5HS2H ribosomal protein
45 7 3.2 114 2 F84266 50S ribosomal prot

ALIGNMENTS

RESULT 1

T36267

probable glutamyl-trna reductase - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C:Accession: T36267

R:Murphy, L.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, June 1999

A:Reference number: Z21576

A:Accession: T36267

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-581 <MUR>

A:Cross-references: EMBL:AL079345; PIDN:CAB45353.1; GSPDB:GN00070; SCOEDB:SC568.17c

A:Experimental source: strain A3(2)

C:Genetics:

A:Gene: SCOEDB:SC568.17c

Query Match

Best Local Similarity 4.1%; Score 9; DB 2; Length 581;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 154 VAAEPAAEA 162

Db 36 VAAEPAAEA 44

RESULT 2

JC1422

nitrate reductase (NADH) (EC 1.7.1.1) - Volvox carteri

C:Species: Volvox carteri

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 03-Jun-2002

C:Accession: JC1422; S22192

R:Gruber, H.; Goetinck, S.D.; Kirk, D.L.; Schmitt, R.

Gene 120, 75-83, 1992

A:Title: The nitrate reductase-encoding gene of Volvox carteri

A:Reference number: JC1422; MUID:93013022; PMID:1398126

A:Accession: JC1422

A:Molecule type: DNA

A:Residues: 1-864 <GR2>

A:Cross-references: EMBL:X64136; NID:q21993; PIDN:CAA45497.1; PID:q21994

A:Note: submitted to the EMBL Data Library, January 1992

C:Genetics:

A:Gene: nita

A:Map position: linkage group IX

A:Introns: 183/3; 234/2; 295/1; 339/1; 372/1; 425/2; 521/3; 593/3; 677/3; 797/2

C:Function:

A:Description: catalyzes the reduction of nitrate to nitrite

C:Superfamily: nitrate reductase (NADH); cytochrome b5 core homology; cytochrome-b5

C:Keywords: dimer; electron transfer; FAD; flavoprotein; heme; iron; metalloprotein;

F;36-425/Domain: molybdopterin-binding domain homology <PCO>

F;497-571/Domain: cytochrome b5 core homology <CB5>

F:613-864/Domain: cytochrome-b5 reductase homology <CBR>
F:139/Binding site: molybdopterin (Cys) (covalent) #status predicted
F:532,555/Binding site: heme iron (His) (axial ligands) #status predicted

Query Match 4.18; Score 9; DB 1; Length 864;
Best Local Similarity 100.0%; Pred. No. 2.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 AAAAPAAP 128

DB 478 AAAAPAAP 486

RESULT 3

R6BVP3

60s acidic ribosomal protein p1-alpha - fission yeast (Schizosaccharomyces pombe)

N:Alternate names: acidic ribosomal protein 3; ribosomal protein Y12e11

C:Species: Schizosaccharomyces pombe

C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 10-Dec-1999

C:Accession: C34715; T40352

R:Beltrame, M.; Bianchi, M.E.

Mol. Cell. Biol. 10, 2341-2348, 1990

A:Title: A gene family for acidic ribosomal proteins in Schizosaccharomyces pombe: two e

A:Reference number: A34715; MUID:90220620; PMID:2325655

A:Accession: C34715

A:Molecule type: DNA

A:Residues: 1-110 <BEL>

A:Cross-references: EMBL:M33139; NID:g173467; PIDN:AAA35336.1; PID:g173468

R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Skelton, J.; Churcher, C.M.

submitted to the EMBL Data Library, March 1997

A:Reference number: Z21922

A:Accession: T40352

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-110 <WOO>

A:Cross-references: EMBL:AL022070; PIDN:CAA17793.1; GSPDB:GN00067; SPDB:SPBC3B9.13c

A:Experimental source: strain 972h-; cosmid c3B9

C:Genetics:

A:Gene: rpa3

A:Map position: 2

A:Insertions: 23/3; 103/3

C:Superfamily: rat acidic ribosomal protein p1

C:Keywords: phosphoprotein; protein biosynthesis; ribosome

Query Match 3.7%; Score 8; DB 1; Length 110;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 119 SAAAPAA 126

DB 64 SAAAPAA 71

RESULT 4

H86500

L7/L12 ribosomal protein [Imported] - Chlamydomophila pneumoniae (strain J138)

C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001

C:Accession: H86500

R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Is

Nucleic Acids Res. 28, 2311-2314, 2000

A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.

A:Reference number: A86491; MUID:20330349; PMID:10871362

A:Accession: H86500

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-129 <STO>

A:Cross-references: GB:BA000008; NID:g8978453; PIDN:BAA98290.1; GSPDB:GN00142

A:Experimental source: strain J138

C:Genetics:

A:Gene: r17

C:Superfamily: Escherichia coli ribosomal protein L12

Query Match

3.7%; Score 8; DB 2; Length 193;

Query Match 3.7%; Score 8; DB 2; Length 129;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 171 EAPVAAEP 178

DB 50 EAPVAAEP 57

RESULT 5

C72122

ribosomal protein L7/L12 CP0695 [Imported] - Chlamydomophila pneumoniae (strains CML029

C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae

C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000

C:Accession: C72122; G81548

R:Kalan, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood,

Nature Genet. 21, 385-389, 1999

A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.

A:Reference number: A72000; MUID:99206606; PMID:10192388

A:Accession: C72122

A:Molecule type: DNA

A:Residues: 1-129 <ARN>

A:Cross-references: GB:AE001593; GB:AE001363; NID:g4376334; PIDN:AAD18233.1; PID:g437

A:Experimental source: strain CML029

R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heideberg, J.F.; White, O.; Hicke

, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzbe

Nucleic Acids Res. 28, 1397-1406, 2000

A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39

A:Reference number: AB1500; MUID:20150255; PMID:10684935

A:Accession: G81548

A:Molecule type: DNA

A:Residues: 1-129 <REA>

A:Cross-references: GB:AE002228; GB:AE002161; NID:g7189606; PIDN:AAF38503.1; PID:g718

A:Experimental source: strain AR39, HL cells

C:Genetics:

A:Gene: r17; CP0695

C:Superfamily: Escherichia coli ribosomal protein L12

Query Match 3.7%; Score 8; DB 2; Length 129;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 171 EAPVAAEP 178

DB 50 EAPVAAEP 57

RESULT 6

H86404

probable lipid transfer protein [Imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: H86404

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, T.H.; Dewar,

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzia

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: AB6141; MUID:21016719; PMID:11130712

A:Accession: H86404

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-193 <STO>

A:Cross-references: GB:AE005172; NID:g11024871; PIDN:AAG26955.1; GSPDB:GN00141

C:Genetics:

A:Map position: 1

Best Local Similarity 100.0%; Pred. No. 7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 AAAAPAAP 127
|||||
Db 20 AAAAPAAP 27

RESULT 7

E87612
cytochrome c, membrane-bound [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: E87612
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
N.; J.; Emolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of *Caulobacter crescentus*.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: E87612
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-228 <STO>
A:Cross-references: GB:AE005673; NID:gl3424561; PIDN:AAK24897.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC2935

Query Match 3.7%; Score 8; DB 2; Length 228;
Best Local Similarity 100.0%; Pred. No. 8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 AAAAPAAP 127
|||||
Db 219 AAAAPAAP 226

RESULT 8

G70645
hypothetical protein Rv3134c - Mycobacterium tuberculosis (strain H37Rv)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: G70645
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: G70645
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-268 <COL>
A:Cross-references: GB:Z83867; GB:AL123456; NID:g3261695; PIDN:CAB06280.1; PID:e290951;
A:Experimental source: strain H37Rv
C:Genetics:
A:Gene: Rv3134c

Query Match 3.7%; Score 8; DB 2; Length 268;
Best Local Similarity 100.0%; Pred. No. 9.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 PSPAEPAT 200
|||||
Db 138 PSPAEPAT 145

RESULT 9

G95887
probable ABC transporter permease protein Smb20381 [imported] - *Sinorhizobium meliloti*
C:Species: *Sinorhizobium meliloti*
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 01-Mar-2002

C:Accession: G95887
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hei
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A:Title: The complete sequence of the 1.683-kb pSymB megaplasmid from the N₂-fixing
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: G95887
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-276 <KUR>
A:Cross-references: GB:AL591985; PIDN:CAC48767.1; PID:gl5140240; GSPDB:GN00167
A:Experimental source: strain 1021, megaplasmid pSymB
Pela, D.; Chain, P.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hub
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lela
heault, P.; Vandebol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh,
A:Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: Smb20381
A:Genome: plasmid
C:Superfamily: spermidine/putrescine transport system permease protein potI

Query Match 3.7%; Score 8; DB 2; Length 276;
Best Local Similarity 100.0%; Pred. No. 9.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 AAEPAAAA 137
|||||
Db 268 AAEPAAAA 275

RESULT 10

T02434
DNA binding protein EREBP-4 - common tobacco
C:Species: *Nicotiana tabacum* (common tobacco)
C:Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 31-Mar-2001
C:Accession: T02434
R:Omme-Takagi, M.; Shinshi, H.
Plant Cell 7, 173-182, 1995
A:Title: Ethylene-inducible DNA binding proteins that interact with an ethylene resp
A:Reference number: Z14671; MUID:95276459; PMID:7756828
A:Accession: T02434
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-291 <OHM>
A:Cross-references: EMBL:D38125; NID:g790361; PIDN:BAA07323.1; PID:gl208497
A:Experimental source: strain BY4; tissue-type leaf
C:Superfamily: Arabidopsis thaliana hypothetical protein T19P19.170

Query Match 3.7%; Score 8; DB 2; Length 291;
Best Local Similarity 100.0%; Pred. No. 9.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 133 PAAAPLPT 140
|||||
Db 247 PAAAPLPT 254

RESULT 11

T48873
electron transfer protein [imported] - *Fraterulia* sp. (strain ANA-18)
C:Species: *Fraterulia* sp.
A:Variety: strain ANA-18
C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jun-2000
C:Accession: T48873
R:Murakami, S.; Takashima, A.; Takemoto, J.; Takenaka, S.; Shinke, R.; Aoki, K.
Gene 226, 189-198, 1999
A:Title: Cloning and sequence analysis of two catechol-degrading gene clusters from t
A:Reference number: Z24833
A:Accession: T48873

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-325 <MUR>
A:Cross-references: EMBL:AB009373; PIDN:BAA75213.1
A:Experimental source: strain ANA-18
C:Superfamily: phthalate dioxygenase reductase; cytochrome-b5 reductase homology; ferredoxin

Query Match 3.7%; Score 8; DB 2; Length 325;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 130 AAEPAAAA 137
Db 223 AAEPAAAA 230

RESULT 12

A87300
TPR domain protein [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: A87300
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.; B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolonitskii, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.; Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: A87300
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-336 <SPT>
A:Cross-references: GB:AE005673; NID:g13421571; PIDN:AAK22397.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC0410

Query Match 3.7%; Score 8; DB 2; Length 336;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 120 AAAAPAAP 127
Db 112 AAAAPAAP 119

RESULT 13

S41958
Pupal cuticle protein precursor - greater wax moth
C:Species: Galleria mellonella (greater wax moth)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 19-May-2000
C:Accession: S41958
R:Kolliberg, U.; Obermeier, B.; Hirsch, H.; Kelber, G.; Wolbert, P.
submitted to the EMBL Data Library, February 1994
A:Description: Expression cloning and characterization of a pupal cuticle protein cDNA
A:Reference number: S41958
A:Accession: S41958
A:Molecule type: mRNA
A:Residues: 1-353 <KOL>
A:Cross-references: EMBL:X77514; NID:g453403; PIDN:CAA54650.1; PID:g453404
F:1-15/Domain: signal sequence #status predicted <SIG>
F:16-353/Product: pupal cuticle protein #status predicted <MAT>

Query Match 3.7%; Score 8; DB 2; Length 353;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 119 SAAAPAA 126
Db 236 SAAAPAA 243

RESULT 14

T46827

phenoxybenzoate dioxygenase [imported] - Acinetobacter lwoffii
C:Species: Acinetobacter lwoffii
C:Date: 17-Mar-2000 #sequence_revision 17-Mar-2000 #text_change 02-Sep-2000
C:Accession: T46827
R:Kim, S.I.; Leem, S.H.; Choi, J.S.; Chung, Y.H.; Kim, S.; Park, Y.M.; Ha, K.S.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z24100
A:Accession: T46827
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-381 <KIM>
A:Cross-references: EMBL:U77659; PIDN:AAC31770.1
A:Experimental source: strain K24
C:Superfamily: phthalate dioxygenase reductase; cytochrome-b5 reductase homology; ferredoxin

Query Match 3.7%; Score 8; DB 2; Length 381;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 130 AAEPAAAA 137
Db 279 AAEPAAAA 286

RESULT 15

JH0633
cellular tumor antigen p53 - golden hamster
N:Alternate names: tumor-suppressor protein p53
C:Species: Mesocricetus auratus (golden hamster)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: JH0633
R:Legros, Y.; McIntyre, P.; Soussi, T.
Gene 112, 247-250, 1992
A:Title: The cDNA cloning and immunological characterization of hamster p53.
A:Reference number: JH0633; MUID:92210007; PMID:1555773
A:Accession: JH0633
A:Molecule type: mRNA
A:Residues: 1-396 <LEG>
A:Cross-references: GB:M75144; NID:g191414; PIDN:AAA37085.1; PID:g191415
A:Experimental source: kidney, strain MP1
C:Genetics:
A:Gene: p53
C:Superfamily: cellular tumor antigen p53
C:Keywords: apoptosis; cell division control; DNA binding; homotetramer; nucleus; phosphatase; zinc finger
F:179,182,241,245/Binding site: zinc (Cys, His, Cys) #status predicted
F:395/Binding site: phosphoryl-RNA (Ser) (covalent) #status predicted

Query Match 3.7%; Score 8; DB 1; Length 396;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 120 AAAAPAAP 127
Db 66 AAAAPAAP 73

Search completed: July 7, 2003, 15:12:56
Job time : 44 secs

DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE 60S acidic ribosomal protein P0.
GN LIPO-A AND LIPO-B.
OS Leishmania infantum.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5671;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-LEM 75;
RX MEDLINE=94088674; PubMed=8264730;
RA Soto M., Requena J.M., Alonso C.;
RT "Isolation, characterization and analysis of the expression of the
RT Leishmania ribosomal P0 protein genes";
RL Mol. Biochem. Parasitol. 61:265-274(1993).
CC -1- FUNCTION: RIBOSOMAL PROTEIN P0 IS THE FUNCTIONAL EQUIVALENT
CC OF E.COLI PROTEIN L10.
CC -1- SUBUNIT: P0 FORMS A PENTAMERIC COMPLEX BY INTERACTION WITH
CC DIMERS OF P1 AND P2.
CC -1- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X72714; CAA51264.1; -;
DR EMBL; X72714; CAA51263.1; -;
DR InterPro; IPR001813; 60S.Ribosomal.
DR InterPro; IPR001790; Ribosomal.L10.
DR Pfam; PF00428; 60S.Ribosomal; 1.
DR Pfam; PF00466; Ribosomal.L10; 1.
KW Ribosomal protein; Phosphorylation.
SQ SEQUENCE 323 AA; 34771 MW; 149677BD7A439D69E CRC64;

Query Match 4.1%; Score 9; DB 1; Length 323;
Best Local Similarity 100.0%; Pred. No. 0.57;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 130 AAEPAAAAP 138
| | | | | | | |
Db 290 AAEPAAAAP 298

RESULT 3
ID HEMI_STRCO STANDARD; PRT; 581 AA.
AC Q9WX15;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Glutamyl-tRNA reductase (EC 1.2.1.-) (GluR).
GN HEMA OR SC03319 OR SCE68.17C.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinobacteriales;
OC Actinomycetales; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1502;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Gobie A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,

RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
CC -1- CATALYTIC ACTIVITY: Glutamyl-tRNA(Glu) + NADPH = glutamate-1-
CC semialdehyde + NADP(+) + tRNA(Glu).
CC -1- PATHWAY: Porphyrin biosynthesis by the C5 pathway; first step.
CC -1- SIMILARITY: BELONGS TO THE GLUTAMYL-TRNA REDUCTASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AL079345; CAB45353.1; -;
DR InterPro; IPR000343; GluTR.
DR InterPro; IPR000594; Thif_domain.
DR Pfam; PF00745; GluTR; 1.
DR TIGRFAMs; TIGR01035; hema; 1.
DR PROSITE; PS00747; GLUTR; 1.
KW Porphyrin biosynthesis; Oxidoreductase; NADP; Complete proteome.
FT DOMAIN 232 416 INSERT.
FT ACT_SITE 50 50 NUCLEOPHILE (BY SIMILARITY).
FT ACT_SITE 99 99 BASE (BY SIMILARITY).
SQ SEQUENCE 581 AA; 60562 MW; D4E256B105AFA037 CRC64;

Query Match 4.1%; Score 9; DB 1; Length 581;
Best Local Similarity 100.0%; Pred. No. 0.93;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 154 VAAEPAAEA 162
| | | | | | | |
Db 36 VAAEPAAEA 44

RESULT 4
ID NIA_VOLCA STANDARD; PRT; 864 AA.
AC P36841;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE Nitrate reductase (EC 1.6.6.1) (NR).
GN NITA.
OS Volvox carteri.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Volvocaceae; Volvox.
OX NCBI_TaxID=3067;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=f. Nagariensis / HK10;
RX MEDLINE=93013022; PubMed=1398126;
RA Gruber H., Goetinck S.D., Kirk D.L., Schmitt R.;
RT "The nitrate reductase-encoding gene of Volvox carteri: map location,
RT sequence and induction kinetics";
RL Gene 120:75-83(1992).
CC -1- FUNCTION: NITRATE REDUCTASE IS A KEY ENZYME INVOLVED IN THE FIRST
CC STEP OF NITRATE ASSIMILATION IN PLANTS, FUNGI AND BACTERIA.
CC -1- CATALYTIC ACTIVITY: NADH + nitrate = NAD(+) + nitrite + H(2)O.
CC -1- COFACTOR: EACH SUBUNIT OF THE ENZYME CONTAINS 1 EQUIVALENT OF FAD,
CC HEME IRON, AND MOLYBDENUM-PTERIN AS PROSTHETIC GROUPS. THE HEME
CC GROUP IS CALLED CYTOCHROME B-557.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- INDUCTION: BY NITRATE.
CC -1- SIMILARITY: TO EUKARYOTIC MOLYBDOPTEIN OXIDOREDUCTASES IN THE
CC N-TERMINAL DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 CYTOCHROME B5 FAMILY, HEME-BINDING DOMAIN.
CC -1- SIMILARITY: TO FAD/NAD-BINDING CYTOCHROME REDUCTASES IN THE
CC C-TERMINAL DOMAIN.
CC -----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

CC EMBL: X64136; CAA45497.1;
CC PIR: JC1422; JC1422.
CC HSSP: P04166; 1B5M.
CC InterPro: IPR001199; Cyt_B5.
CC InterPro: IPR001834; Cyt_B5_reductase.
CC InterPro: IPR000572; Euk_Mb_oxred.
CC InterPro: IPR005066; Mo-co_dimer.
CC InterPro: IPR001433; Oxred_FAD/NAD(P).
CC Pfam: PF00173; heme_1; 1.
CC Pfam: PF00174; Oxidored_molyb; 1.
CC Pfam: PF00175; NAD_binding; 1.
CC Pfam: PF00970; FAD_binding_6; 1.
CC Pfam: PF03404; Mo-co_dimer; 1.
CC PRINTS: PR00406; CYTB5RDTASE.
CC PRINTS: PR00363; CYTOCHROME5.
CC PRINTS: PR00407; EUMOPTERIN.
CC ProDom: PD000612; Cyt_B5; 1.
CC PROSITE: PS00191; CYTOCHROME_B5_1; 1.
CC PROSITE: PS02555; CYTOCHROME_B5_2; 1.
CC PROSITE: PS00559; MOLYBDOPTERIN_EUK; 1.
CC Oxidoreductase: Flavoprotein; FAD; NAD; Heme; Molybdenum;
KW Nitrate assimilation.
FT METAL 139 139 MOLYBDENUM-PTERIN (POTENTIAL).
FT METAL 193 193 MOLYBDENUM-PTERIN (POTENTIAL).
FT DISULFID 376 376 INTERCHAIN (POTENTIAL).
FT BINDING 532 532 HEME LIGAND (BY SIMILARITY).
FT BINDING 555 555 HEME LIGAND (BY SIMILARITY).
SQ SEQUENCE 864 AA; 96402 MW; 499529652CDDDC1C7 CRC64;

Query Match 4.1%; Score 9; DB 1; Length 864;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 AAAAPAAP 128
DB 478 AAAAPAAP 486

RESULT 5
RLAL LEIPE STANDARD; PRT; 107 AA.
AC 046313.
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE 60S acidic ribosomal protein p1.
OS Leishmania peruviana.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5681;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MHOM/PE/84/LC26;
RA de Los Santos M., Carrillo C., Panebra A., Montoya Y.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE ELONGATION STEP OF
CC PROTEIN SYNTHESIS.
CC -!- SUBUNIT: P1 AND P2 EXIST AS DIMERS AT THE LARGE RIBOSOMAL
CC SUBUNIT.
CC -!- SIMILARITY: BELONGS TO THE L12P FAMILY OF RIBOSOMAL PROTEINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```

CC EMBL: AF045249; AAC02701.1;
CC InterPro: IPR001813; 60s_ribosomal.
CC Pfam: PF00428; 60s_ribosomal; 1.
CC Ribosomal protein.
SQ SEQUENCE 107 AA; 10896 MW; 9117DEC5E490A071 CRC64;

Query Match 3.7%; Score 8; DB 1; Length 107;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 119 SAAAPAA 126
DB 71 SAAAPAA 78

RESULT 6
RLA3 SCHPO STANDARD; PRT; 110 AA.
ID RLA3_SCHPO
AC P17477;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 60S acidic ribosomal protein P1-alpha 3 (A3).
GN RPA3 OR SPBC3B9.13C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=90220620; PubMed=2325655;
RA Beltrame M., Bianchi M.E.;
RT "A gene family for acidic ribosomal proteins in Schizosaccharomycetes
RL pombe: two essential and two nonessential genes.";
RL Mol. Cell. Biol. 10:2341-2348(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart G., Aert R., Robben J., Grymonprez B.,
RA Welljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., LeFrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Carrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Sipakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomycetes pombe.";
RL Nature 415:871-880(2002).
CC -!- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE ELONGATION STEP OF
CC PROTEIN SYNTHESIS.
CC -!- SUBUNIT: P1 AND P2 EXIST AS DIMERS AT THE LARGE RIBOSOMAL

```

CC SUBUNIT.
CC -1- MISCELLANEOUS: YEASTS CONTAIN 4 INDIVIDUAL SMALL RIBOSOMAL A
CC PROTEINS (RPA) WHICH CAN BE CLASSIFIED INTO TWO COUPLES OF SIMILAR
CC BUT NOT IDENTICAL SEQUENCES. EACH COUPLE IS DISTINCTLY RELATED TO
CC ONE OF THE TWO A PROTEINS PRESENT IN MULTICELLULAR ORGANISMS.
CC -1- MISCELLANEOUS: RPA3 AND RPA4 ARE ESSENTIAL FOR CELL SURVIVAL,
CC WHEREAS RPA1 AND RPA2 ARE NOT.
CC -1- SIMILARITY: BELONGS TO THE L12P FAMILY OF RIBOSOMAL PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M33139; AAA35336.1; -
CC EMBL; AL022070; CAAL17793.1; -
CC PIR; C34715; R6BYP3.
CC InterPro: IPR001813; 60s_ribosomal.
CC Pfam; PF00428; 60s_ribosomal; 1
CC Ribosomal protein; Phosphorylation; Multigene family.
CC KW
CC SEQUENCE 110 AA; 11171 MW; 1734AC9779F5A891 CRC64;
CC
CC Query Match 3.7%; Score 8; DB 1; Length 110;
CC Best Local Similarity 100.0%; Pred. No. 1.9;
CC Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 119 SAAAPAA 126
CC Db 64 SAAAPAA 71
CC
CC RESULT 7
CC RL7_CHLPN STANDARD; PRT; 128 AA.
CC ID RL7_CHLPN
CC AC Q949A1; Q9JQ70;
CC DT 30-MAY-2000 (Rel. 39, Created)
CC DT 30-MAY-2000 (Rel. 39, Last sequence update)
CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
CC DE 50S ribosomal protein L7/L12.
CC GN RPL7 OR RL7 OR CPN0080 OR CP0695.
CC OS Chlamydia pneumoniae (Chlamydia pneumoniae).
CC OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
CC NCBI_TaxID=83558;
CC [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN-CWL029;
CC RX MEDLINE=99206606; PubMed=10192388;
CC RA Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
CC Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
CC "Comparative genomes of Chlamydia pneumoniae and C. trachomatis";
CC Nat. Genet. 21:385-389(1999).
CC [2]
CC RN
CC RP SEQUENCE FROM N.A.
CC RC STRAIN-AR39;
CC RX MEDLINE=20150255; PubMed=10684935;
CC RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
CC White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
CC Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
CC Gwinn M., Nelson W., Deboy R., Kolonay J., McClarty G., Salzberg S.L.,
CC Eisen J., Fraser C.M.;
CC "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
CC pneumoniae AR39";
CC RL Nucleic Acids Res. 28:1397-1406(2000).
CC [3]
CC RN
CC RP SEQUENCE FROM N.A.
CC RX STRAIN-J138;
CC RX MEDLINE=20330349; PubMed=10871362;
CC RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
CC Shiba T., Ishi K., Hattori M., Kuhara S., Nakazawa T.;
CC "Comparison of whole genome sequences of Chlamydia pneumoniae J138

RT from Japan and CWL029 from USA.";
RL Nucleic Acids Res. 28:2311-2314(2000).
CC -1- FUNCTION: SEEMS TO BE THE BINDING SITE FOR SEVERAL OF THE FACTORS
CC INVOLVED IN PROTEIN SYNTHESIS AND APPEARS TO BE ESSENTIAL FOR
CC ACCURATE TRANSLATION (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE L12P FAMILY OF RIBOSOMAL PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE001593; AAD18233.1; -
CC EMBL; AE002228; AAF38503.1; -
CC EMBL; AF002545; BAA98290.1; -
CC HSSP; P02392; ICTF.
CC PHCI-2DPAGE; Q9Z9A1; -
CC TIGR; CP0695; -
CC InterPro: IPR000206; Ribosomal_L12.
CC Pfam; PF00542; Ribosomal_L12; 1.
CC Prodom; PD001326; Ribosomal_L12; 1.
CC DR TIGRFAMS; TIGR00855; L12; 1.
CC KW Ribosomal protein; Complete proteome.
CC FT INIT_MET 0 0 BY SIMILARITY.
CC SQ SEQUENCE 128 AA; 13461 MW; 4E2F17A85B057CC CRC64;
CC
CC Query Match 3.7%; Score 8; DB 1; Length 128;
CC Best Local Similarity 100.0%; Pred. No. 2.2;
CC Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 171 EAPVAAEP 178
CC Db 49 EAPVAAEP 56
CC
CC RESULT 8
CC CUP5_GALME STANDARD; PRT; 353 AA.
CC ID CUP5_GALME
CC AC Q24998;
CC DT 15-DEC-1998 (Rel. 37, Created)
CC DT 15-DEC-1998 (Rel. 37, Last sequence update)
CC DT 15-JUL-1999 (Rel. 38, Last annotation update)
CC DE Pupal cuticle protein PCP52 precursor (GMPCP52).
CC GN PCP52.
CC OS Galleria mellonella (Wax moth).
CC OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
CC Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata;
CC Ditrysia; Pyraloidea; Pyralidae; Galleriinae; Galleria.
CC NCBI_TaxID=7137;
CC [1]
CC RN
CC RP SEQUENCE FROM N.A.
CC RC TISSUE-Epidermis;
CC RX MEDLINE=95291282; PubMed=7773255;
CC RA Kolberg U., Obermaier B., Hirsch H., Kelber G., Wolbert P.;
CC "Expression cloning and characterization of a pupal cuticle protein
CC cDNA of Galleria mellonella L.";
CC Insect Biochem. Mol. Biol. 25:355-363(1995).
CC -1- FUNCTION: COMPONENT OF THE CUTICLE OF THE PUPA OF GALLERIA
CC MELLONELLA.
CC -1- DEVELOPMENTAL STAGE: MAXIMAL EXPRESSION DURING THE FIRST DAY AFTER
CC PUPAL ECDYSIS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC

DR EMBL: X77514; CAA54650.1; -.
 KW Structural protein; Cuticle; Signal.
 FT SIGNAL 1 15 POTENTIAL.
 FT CHAIN 16 353 PUPAL CUTICLE PROTEIN PCP52.
 FT DOMAIN 235 243 POLY-ALA.
 SQ SEQUENCE 353 AA; 35818 MW; C3A35B04EC049172 CRC64;
 Query Match 3.7%; Score 8; DB 1; Length 353;
 Best Local Similarity 100.0%; Pred. No. 5.2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 119 SAAAPAA 126
 Db 236 SAAAPAA 243
 RESULT 9
 P53_MESAU
 ID P53_MESAU STANDARD; PRT; 396 AA.
 AC Q00366; P97276;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cellular tumor antigen p53 (Tumor suppressor p53).
 GN TP53.
 OS Mesocricetus auratus (Golden hamster).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
 OC Mesocricetus.
 OX NCBI_TaxID=10036;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Syrian; TISSUE=Kidney;
 RX MEDLINE=92210007; PubMed=1555773;
 RA Legros Y., McIntyre P., Soussi T.;
 RT "The cDNA cloning and immunological characterization of hamster p53";
 RL Gene 112:247-250(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Hou E.W., Wiseman R.;
 RL Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Acts as a tumor suppressor in many tumor types; Induces growth arrest or apoptosis depending on the physiological circumstances and cell type. Involved in cell cycle regulation as a trans-activator that acts to negatively regulate cell division by controlling a set of genes required for this process. One of the activated genes is an inhibitor of cyclin-dependent kinases. Apoptosis induction seems to be mediated either by stimulation of BAX and FAS antigen expression, or by repression of Bcl-2 expression.
 CC -!- SUBUNIT: Binds DNA as a homotetramer (By similarity).
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- DISEASE: P53 IS FOUND IN INCREASED AMOUNTS IN A WIDE VARIETY OF TRANSFORMED CELLS. P53 IS FREQUENTLY MUTATED OR INACTIVATED IN MANY TYPES OF CANCER.
 CC -!- SIMILARITY: BELONGS TO THE P53 FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
 CC
 CC EMBL: M75144; AAA37085.1; -.
 CC EMBL: U07182; AAB41344.1; -.
 CC PIR: JH0633; JH0633.
 CC HSSP: P04637; 1TUP.
 CC InterPro: IPR002117; P53.
 CC Pfam: PF00870; P53; 1.
 CC PRINTS: PR00386; P53SUPPRESSOR.
 CC ProDom: PD002681; P53; 1.

DR PROSITE; PS00348; P53; 1.
 KW Anti-oncogene; DNA-binding; Transcription regulation; Activator;
 KW Nuclear protein; Phosphorylation; Apoptosis.
 FT DOMAIN 1 45 TRANSCRIPTION ACTIVATION (ACIDIC).
 FT DNA_BIND 105 295 BY SIMILARITY.
 FT DOMAIN 328 359 OLIGOMERIZATION.
 FT DOMAIN 371 390 BASIC (REPRESSION OF DNA-BINDING).
 FT DOMAIN 314 326 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT MOD_RES 395 395 PHOSPHORYLATION (BY SIMILARITY).
 FT CONFLICT 188 188 G -> S (IN REF. 2).
 SQ SEQUENCE 396 AA; 43631 MW; 906EF0256809BE3 CRC64;
 Query Match 3.7%; Score 8; DB 1; Length 396;
 Best Local Similarity 100.0%; Pred. No. 5.7;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 120 AAAAPAAP 127
 Db 66 AAAAPAAP 73
 RESULT 10
 ATP2_CHLRE
 ID ATP2_CHLRE STANDARD; PRT; 574 AA.
 AC P38482;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE ATP synthase beta chain, mitochondrial precursor (EC 3.6.3.14).
 GN ATP2.
 OS Chlamydomonas reinhardtii.
 OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
 OC Chlamydomonadaceae; Chlamydomonas.
 OX NCBI_TaxID=3055;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92353386; PubMed=1386535;
 RA Franzen L.-G., Falk G.;
 RT "Nucleotide sequence of cDNA clones encoding the beta subunit of mitochondrial ATP synthase from the green alga Chlamydomonas reinhardtii: the precursor protein encoded by the cDNA contains both an N-terminal presequence and a C-terminal extension.";
 RL Plant Mol. Biol. 19:771-780(1992).
 RN [2]
 RP STRUCTURE BY NMR OF 1-26.
 RX MEDLINE=96326639; PubMed=8706917;
 RA Lancelin J.-M., Gans P., Bouchayer E., Bally I., Arian G.J., Jacquot J.-P.;
 RT "NMR structures of a mitochondrial transit peptide from the green alga Chlamydomonas reinhardtii.";
 RL FEBS Lett. 391:203-208(1996).
 CC -!- FUNCTION: PRODUCES ATP FROM ADP IN THE PRESENCE OF A PROTON GRADIENT ACROSS THE MEMBRANE. THE BETA CHAIN IS THE CATALYTIC SUBUNIT.
 CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate + H(+) (Out).
 CC -!- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0) HAS THREE MAIN SUBUNITS: A, B AND C.
 CC -!- SUBCELLULAR LOCATION: Mitochondrial.
 CC -!- SIMILARITY: BELONGS TO THE ATPASE ALPHA/BETA CHAINS FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X61624; CAA43808.1; -.
 CC PIR: S23530; S23530.

```

DR HSSP: P00829; 1BMF.
DR InterPro: IPR000793; ATPase_a/bc.
DR InterPro: IPR004100; ATPase_a/bn.
DR InterPro: IPR000194; ATPase_a/bcentre.
DR Pfam: PF00006; ATP-synt_ab; 1.
DR Pfam: PF00306; ATP-synt_ab.C; 1.
DR Pfam: PF02874; ATP-synt_ab.N; 1.
DR TIGRFAMs: TIGR01039; atpD; 1.
DR PROSITE: PS00152; ATPASE_ALPHA_BETA; 1.
KW ATP synthesis; CF(1); Hydrogen ion transport;
KW Hydrolase; ATP-binding; Mitochondrion; Transit peptide.
FT TRANSIT 1 26 MITOCHONDRION (POTENTIAL).
FT CHAIN 27 574 ATP SYNTHASE BETA CHAIN.
FT NP_BIND 183 190 ATP (BY SIMILARITY).
SQ SEQUENCE 574 AA; 61821 MW; 22B3C6C6D18FBCFE CRC64;

Query Match 3.7%; Score 8; DB 1; Length 574;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 130 AAEPAAAA 137
Db 26 AAEPAAAA 33

RESULT 11
MSPA_TREMA STANDARD; PRT; 575 AA.
AC Q92413;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Major outer membrane protein mspa precursor (Major sheath protein).
GN MSPA.
OS Treponema maltophilum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=51160;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 20-34.
RC STRAIN=ATCC 51939 / BR;
RX MEDLINE=99121045; PubMed=9922270;
RA Heuner K., Choi B.-K., Schade R., Moter A., Otto A., Goebel U.B.;
RT "Cloning and characterization of a gene (mspa) encoding the major
RT sheath protein of Treponema maltophilum ATCC 51939(T).";
RL J. Bacteriol. 181:1023-1029(1999).
RN [2]
RP SUBCELLULAR LOCALIZATION.
RX PubMed=11313133;
RA Heuner K., Meltzer U., Choi B.-K., Goebel U.B.;
RT "Outer sheath associated proteins of the oral spirochete Treponema
RT maltophilum.";
RL FEMS Microbiol. Lett. 197:187-193(2001).
CC -1- FUNCTION: Major component of the outer membrane sheath.
CC -1- SUBCELLULAR LOCATION: Outer membrane-associated.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: Y17800; CAA76862.1; -.
CC KW Outer membrane; Signal.
CC FT SIGNAL 1 19
CC FT CHAIN 20 575 MAJOR OUTER MEMBRANE PROTEIN MSPA.
CC SQ SEQUENCE 575 AA; 62250 MW; C666B188F5DE16A0 CRC64;

Query Match 3.7%; Score 8; DB 1; Length 575;
Best Local Similarity 100.0%; Pred. No. 7.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DR HSSP: P00829; 1BMF.
DR InterPro: IPR000793; ATPase_a/bc.
DR InterPro: IPR004100; ATPase_a/bn.
DR InterPro: IPR000194; ATPase_a/bcentre.
DR Pfam: PF00006; ATP-synt_ab; 1.
DR Pfam: PF00306; ATP-synt_ab.C; 1.
DR Pfam: PF02874; ATP-synt_ab.N; 1.
DR TIGRFAMs: TIGR01039; atpD; 1.
DR PROSITE: PS00152; ATPASE_ALPHA_BETA; 1.
KW ATP synthesis; CF(1); Hydrogen ion transport;
KW Hydrolase; ATP-binding; Mitochondrion; Transit peptide.
FT TRANSIT 1 26 MITOCHONDRION (POTENTIAL).
FT CHAIN 27 574 ATP SYNTHASE BETA CHAIN.
FT NP_BIND 183 190 ATP (BY SIMILARITY).
SQ SEQUENCE 574 AA; 61821 MW; 22B3C6C6D18FBCFE CRC64;

Query Match 3.7%; Score 8; DB 1; Length 574;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 130 AAEPAAAA 137
Db 26 AAEPAAAA 33

RESULT 11
MSPA_TREMA STANDARD; PRT; 575 AA.
AC Q92413;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Major outer membrane protein mspa precursor (Major sheath protein).
GN MSPA.
OS Treponema maltophilum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=51160;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 20-34.
RC STRAIN=ATCC 51939 / BR;
RX MEDLINE=99121045; PubMed=9922270;
RA Heuner K., Choi B.-K., Schade R., Moter A., Otto A., Goebel U.B.;
RT "Cloning and characterization of a gene (mspa) encoding the major
RT sheath protein of Treponema maltophilum ATCC 51939(T).";
RL J. Bacteriol. 181:1023-1029(1999).
RN [2]
RP SUBCELLULAR LOCALIZATION.
RX PubMed=11313133;
RA Heuner K., Meltzer U., Choi B.-K., Goebel U.B.;
RT "Outer sheath associated proteins of the oral spirochete Treponema
RT maltophilum.";
RL FEMS Microbiol. Lett. 197:187-193(2001).
CC -1- FUNCTION: Major component of the outer membrane sheath.
CC -1- SUBCELLULAR LOCATION: Outer membrane-associated.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: Y17800; CAA76862.1; -.
CC KW Outer membrane; Signal.
CC FT SIGNAL 1 19
CC FT CHAIN 20 575 MAJOR OUTER MEMBRANE PROTEIN MSPA.
CC SQ SEQUENCE 575 AA; 62250 MW; C666B188F5DE16A0 CRC64;

Query Match 3.7%; Score 8; DB 1; Length 575;
Best Local Similarity 100.0%; Pred. No. 7.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 AAEPAAEA 162
DB 19 AAEPAAEA 26

RESULT 12
ECR_HELVI STANDARD; PRT; 576 AA.
AC O18473;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ecdysone receptor (Ecdysteroid receptor) (20-hydroxy-ecdysone
DE receptor) (20E receptor) (HVECR).
GN ECR OR NR1H1.
OS Heliothis virescens (Noctuid moth) (Owlet moth).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata;
OC Ditrysia; Noctuoidea; Noctuidae; Heliothinae; Heliothlis.
OX NCBI_TaxID=7102;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRC;
RX MEDLINE=99457718; PubMed=10528411;
RA Martinez A., Scanlon D., Gross B., Perara S.C., Palli S.R.,
RA Greenland A.J., Windass J., Pongs O., Broad P., Jepson I.;
RT "Transcriptional activation of the cloned Heliothis virescens
RT (Lepidoptera) ecdysone receptor (HVECR) by muristeroneA.";
RL Insect Biochem. Mol. Biol. 29:915-930(1999).
CC -1- FUNCTION: RECEPTOR FOR ECDYSONE. BINDS TO ECDYSONE RESPONSE
CC ELEMENTS (ECRES) (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
CC -1- NR1 SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: Y09009; CAA70212.1; -.
CC DR HSSP: P20393; IAGY.
CC DR InterPro: IPR000536; Hormone_rec_lig.
CC DR InterPro: IPR001628; Znf_C4steroid.
CC DR Pfam: PF00104; hormone_rec; 1.
CC DR Pfam: PF00105; zf-C4; 1.
CC DR PRINTS: PRO0047; STROIDFINGER.
CC DR PRODOM: PD000035; Znf_C4steroid; 1.
CC DR SMART: SM00430; HOLI; 1.
CC DR SMART: SM00399; Znf_C4; 1.
CC DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
CC KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;
CC KW Zinc-finger.
CC FT DOMAIN 1 162 MODULATING (POTENTIAL).
CC FT DNA_BIND 163 228 NUCLEAR RECEPTOR-TYPE.
CC FT ZN_FING 163 183 C4-TYPE.
CC FT ZN_FING 199 223 C4-TYPE.
CC FT DOMAIN 326 545 HORMONE-BINDING (POTENTIAL).
CC SQ SEQUENCE 576 AA; 64638 MW; D13EF787BF263A8 CRC64;

Query Match 3.7%; Score 8; DB 1; Length 576;
Best Local Similarity 100.0%; Pred. No. 7.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 140 TATPVAEE 147
DB 552 TATPVAEE 559

RESULT 13

```


DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Neurogenic locus notch homolog protein 2 precursor (Notch 2) (Notch
 DE B).
 DE GN
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Thymus;
 RA Hamada Y., Higuchi M., Tsujimoto Y.;
 RT "Complete amino acid sequence and multiform transcripts encoded by a
 RT single copy of mouse Notch2 gene.";
 RL Submitted (Jul-1994) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RN SEQUENCE OF 316-1518 FROM N.A.
 RC STRAIN=C57BL/6 X CBA; TISSUE=Embryo;
 EX MEDLINE=93178563; PubMed=8440332;
 RA Lardelli M., Lendahl U.;
 RT "Notch A and Notch B-two mouse Notch homologues coexpressed in a
 RT wide variety of tissues.";
 RL Exp. Cell Res. 204:364-372(1993).
 RN [3]
 RN SEQUENCE OF 1765-2153 FROM N.A.
 EX MEDLINE=97075110; PubMed=8917536;
 RA Milner L.A., Bigas A., Kopan R., Brashem-Stein C., Bernstein I.D.,
 RA Martin D.I.;
 RT "Inhibition of granulocytic differentiation by mNotch1.";
 RL Proc. Natl. Acad. Sci. U.S.A. 93:13014-13019(1996).
 RN [4]
 RN FUNCTION.
 EX MEDLINE=99396706; PubMed=10393120;
 RA Hamada Y., Kadohawa Y., Okabe M., Ikawa M., Coleman J.R.,
 RA Tsujimoto Y.;
 RT "Mutation in ankyrin repeats of the mouse Notch2 gene induces early
 RT embryonic lethality.";
 RL Development 126:3415-3424(1999).
 RN [5]
 RN DEVELOPMENTAL STAGE, AND ALTERNATIVE SPLICING.
 EX MEDLINE=9533893; PubMed=7609614;
 RA Higuchi M., Kiyama H., Hayakawa T., Hamada Y., Tsujimoto Y.;
 RT "Differential expression of Notch1 and Notch2 in developing and adult
 RT mouse brain.";
 RL Brain Res. Mol. Brain Res. 29:263-272(1995).
 RN [6]
 RN POST-TRANSLATIONAL PROCESSING, AND MUTAGENESIS OF MET-1699.
 EX MEDLINE=21523956; PubMed=11518718;
 RA Saxena M.T., Schroeter E.H., Mumm J.S., Kopan R.;
 RT "Murine notch homologs (N1-4) undergo presenilin-dependent
 RT proteolysis.";
 RL J. Biol. Chem. 276:40268-40273(2001).
 RN [7]
 RN POST-TRANSLATIONAL PROCESSING, AND MUTAGENESIS OF MET-1699.
 EX MEDLINE=21374376; PubMed=11459941;
 RA Mizutani T., Taniguchi Y., Aoki T., Hashimoto N., Honjo T.;
 RT "Conservation of the biochemical mechanisms of signal transduction
 RT among mammalian Notch family members.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9026-9031(2001).
 CC -1- FUNCTION: Functions as a receptor for membrane-bound ligands
 CC Jagged1, Jagged2 and Delta to regulate cell-fate determination.
 CC Upon ligand activation through the released notch intracellular
 CC domain (NICD) it forms a transcriptional activator complex with
 CC Rap-J kappa and activates genes of the enhancer of split locus.
 CC Affects the implementation of differentiation, proliferation and
 CC apoptotic programs (By similarity). May play an essential role in
 CC postimplantation development, probably in some aspect of cell
 CC specification and/or differentiation.
 CC -1- SUBUNIT: Heterodimer of a C-terminal fragment N(TM) and a N-
 CC terminal fragment N(EC) which are probably linked by disulfide
 CC bonds.

CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Following
 CC proteolytical processing NICD is translocated to the nucleus.
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; may be
 CC produced by alternative splicing.
 CC -1- TISSUE SPECIFICITY: Expressed in the brain, liver, kidney,
 CC neuroepithelia, somites, optic vesicles and branchial arches, but
 CC not heart.
 CC -1- DEVELOPMENTAL STAGE: Expressed in the embryonic ventricular zone,
 CC the postnatal ependymal cells, and the choroid plexus throughout
 CC embryonic and postnatal development.
 CC -1- PTM: Synthesized in the endoplasmic reticulum as an inactive form
 CC which is proteolytically cleaved by a furin-like convertase in the
 CC trans-Golgi network before it reaches the plasma membrane to yield
 CC an active, ligand-accessible form. Cleavage results in a C-
 CC terminal fragment N(TM) and a N-terminal fragment N(EC). Following
 CC ligand binding, it is cleaved by TNF-alpha converting enzyme
 CC (TACE) to yield a membrane-associated intermediate fragment called
 CC notch extracellular truncation (NEXT). This fragment is then
 CC cleaved by presenilin dependent gamma-secretase to release a
 CC notch-derived peptide containing the intracellular domain (NICD)
 CC from the membrane.
 CC -1- PTM: Phosphorylated.
 CC -1- SIMILARITY: BELONGS TO THE NOTCH FAMILY.
 CC -1- SIMILARITY: CONTAINS 34.5 EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 2 LIN/NOTCH REPEATS.
 CC -1- SIMILARITY: CONTAINS 6 ANK REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; D32210; BAA22094.1; -
 CC EMBL; X68279; CA448340.1; -
 CC EMBL; U31881; AAC52924.1; -
 CC HSSP; F16109; 1FSB.
 CC MGD; MGI:97364; Notch2.
 CC InterPro: IPR002110; ANK.
 CC InterPro: IPR000152; Asx_hydroxyl.
 CC InterPro: IPR000561; EGF-like.
 CC InterPro: IPR000742; EGF_2.
 CC InterPro: IPR001881; EGF_Ca.
 CC InterPro: IPR001438; EGF_II.
 CC InterPro: IPR000800; Notch.
 CC Pfam; PF00008; EGF; 35.
 CC Pfam; PF00023; ank; 6.
 CC Pfam; PF00066; notch; 2.
 CC PRINTS; PR00010; EGFBL00D.
 CC PRINTS; PR01452; NOTCH.
 CC SMART; SM00248; ANK; 4.
 CC SMART; SM00179; EGF_CA; 22.
 CC SMART; SM00001; EGF_like; 12.
 CC SMART; SM00004; NL; 3.
 CC PROSITE; PS50088; ANK_REPEAT; 4.
 CC PROSITE; PS50297; ANK_REPEAT_REGION; 1.
 CC PROSITE; PS00010; ASX_HYDROXYL; 22.
 CC PROSITE; PS00022; EGF_1; 33.
 CC PROSITE; PS01186; EGF_2; 27.
 CC PROSITE; PS01187; EGF_CA; 22.
 CC Receptor; Transcription regulation; Activator; Differentiation;
 CC Developmental protein; Repeat; ANK repeat; EGF-like domain;
 CC Transmembrane; Glycoprotein; Signal; Phosphorylation;
 CC Alternative splicing.
 CC SIGNAL 1 25
 CC CHAIN 26 2470
 CC CHAIN 1666 2470
 CC CHAIN 1697 2470
 CC DOMAIN 26 1677
 CC TRANSMEM 1678 1698
 CC DOMAIN 1699 2470
 CC POTENTIAL.
 CC NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN 2.
 CC NOTCH EXTRACELLULAR TRUNCATION.
 CC NOTCH INTRACELLULAR DOMAIN.
 CC EXTRACELLULAR (POTENTIAL).
 CC POTENTIAL.
 CC CYTOPLASMIC (POTENTIAL).

```

FT DOMAIN 26 63 EGF-LIKE 1.
FT DOMAIN 64 102 EGF-LIKE 2.
FT DOMAIN 105 143 EGF-LIKE 3.
FT DOMAIN 144 180 EGF-LIKE 4.
FT DOMAIN 182 219 EGF-LIKE 5.
FT DOMAIN 221 256 EGF-LIKE 6 (INCOMPLETE).
FT DOMAIN 258 294 EGF-LIKE 7.
FT DOMAIN 296 334 EGF-LIKE 8.
FT DOMAIN 336 372 EGF-LIKE 9.
FT DOMAIN 373 411 EGF-LIKE 10.
FT DOMAIN 413 452 EGF-LIKE 11.
FT DOMAIN 454 490 EGF-LIKE 12.
FT DOMAIN 492 528 EGF-LIKE 13.
FT DOMAIN 530 566 EGF-LIKE 14.
FT DOMAIN 568 603 EGF-LIKE 15.
FT DOMAIN 605 641 EGF-LIKE 16.
FT DOMAIN 643 678 EGF-LIKE 17.
FT DOMAIN 680 716 EGF-LIKE 18.
FT DOMAIN 718 753 EGF-LIKE 19.
FT DOMAIN 755 791 EGF-LIKE 20.
FT DOMAIN 793 829 EGF-LIKE 21.
FT DOMAIN 831 869 EGF-LIKE 22.
FT DOMAIN 871 907 EGF-LIKE 23.
FT DOMAIN 909 945 EGF-LIKE 24.
FT DOMAIN 947 983 EGF-LIKE 25.
FT DOMAIN 985 1021 EGF-LIKE 26.
FT DOMAIN 1023 1059 EGF-LIKE 27.
FT DOMAIN 1061 1097 EGF-LIKE 28.
FT DOMAIN 1099 1145 EGF-LIKE 29.
FT DOMAIN 1147 1183 EGF-LIKE 30.
FT DOMAIN 1185 1221 EGF-LIKE 31.
FT DOMAIN 1223 1260 EGF-LIKE 32.
FT DOMAIN 1262 1300 EGF-LIKE 33.
FT DOMAIN 1302 1345 EGF-LIKE 34.
FT DOMAIN 1372 1410 EGF-LIKE 35.
FT REPEAT 1418 1454 LIN/NOTCH 1.
FT REPEAT 1501 1533 LIN/NOTCH 2.
FT REPEAT 1825 1869 ANK 1.
FT REPEAT 1874 1903 ANK 2.
FT REPEAT 1907 1937 ANK 3.
FT REPEAT 1941 1970 ANK 4.
FT REPEAT 1974 2003 ANK 5.
FT REPEAT 2007 2036 ANK 6.
FT DOMAIN 1645 1648 POLY-ALA.
FT DOMAIN 1992 1995 POLY-LEU.
FT DOMAIN 2183 2189 POLY-ALA.
FT DOMAIN 2425 2428 POLY-SER.
FT DOMAIN 2445 2450 POLY-GLY.
FT DISULFID 28 41 BY SIMILARITY.
FT DISULFID 35 51 BY SIMILARITY.
FT DISULFID 53 62 BY SIMILARITY.
FT DISULFID 68 79 BY SIMILARITY.

Query Match 3.7%; Score 8; DB 1; Length 2470;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 120 AAAAPAP 127
Db 2183 AAAAPAP 2190
|||||||

RESULT 15
ANPA_PSEAM STANDARD; PRT; 82 AA.
AC P04002;
DT 23-OCT-1986 (Rel. 02, Created)
DT 23-OCT-1986 (Rel. 02, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Antifreeze protein A/B precursor.
OS Pseudopleuronectes americanus (Winter flounder) (Pleuronectes americanus).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

```

```

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
OC Pleuronectoidei; Pleuronectidae; Pseudopleuronectes.
OX NCBI_TaxID=8265;
RN [1]
RX SEQUENCE FROM N.A. (PROTEIN A).
RX MEDLINE=82197490; PubMed=6952188;
RA Davies P.L., Roach A.H., Hew C.-L.;
RT "DNA sequence coding for an antifreeze protein precursor from winter flounder";
RL Proc. Natl. Acad. Sci. U.S.A. 79:335-339(1982).
RN [2]
RX SEQUENCE FROM N.A. (PROTEIN A).
RX MEDLINE=88259336; PubMed=3133486;
RA Scott G.K., Davies P.L., Kao M.H., Fletcher G.L.;
RT "Differential amplification of antifreeze protein genes in the pleuronectinae";
RL J. Mol. Evol. 27:29-35(1988).
RN [3]
RX SEQUENCE FROM N.A. (PROTEIN B).
RX MEDLINE=84264559; PubMed=6086629;
RA Davies P.L., Hough C., Scott G.K., Ng N., White B.N., Hew C.-L.;
RT "Antifreeze protein genes of the winter flounder";
RL J. Biol. Chem. 259:9241-9247(1984).
RN [4]
RX SEQUENCE FROM N.A.
RX MEDLINE=92209995; PubMed=1555765;
RA Davies P.L.;
RT "Conservation of antifreeze protein-encoding genes in tandem repeats";
RL Gene 112:163-170(1992).
RN [5]
RX 3D-STRUCTURE MODELING OF 45-81.
RX MEDLINE=92148833; PubMed=1738160;
RA Chou K.-C.;
RT "Energy-optimized structure of antifreeze protein and its binding mechanism";
RL J. Mol. Biol. 223:509-517(1992).
RN [6]
RX X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS) OF 45-81.
RX MEDLINE=95281060; PubMed=7760940;
RA Siccheri F., Yang D.S.C.;
RT "Ice-binding structure and mechanism of an antifreeze protein from winter flounder";
RL Nature 375:427-431(1995).
CC -|- FUNCTION: ANTIFREEZE PROTEINS LOWER THE BLOOD FREEZING POINT.
CC -|- MISCELLANEOUS: THE SEQUENCE SHOWN IS THAT OF PROTEIN A.
CC -|- SIMILARITY: BELONGS TO THE TYPE-I AFP FAMILY. TYPE I AFP ARE ALANINE-RICH, AMPHIPHILIC AND ALPHA-HELICAL.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L00138; AAB59964.1;
CC EMBL; L29178; AAB59964.1; JOINED.
CC EMBL; M62414; AAA49469.1;
CC EMBL; X07506; CAA30389.1;
CC EMBL; M62416; AAA49471.1;
CC EMBL; M62417; AAA49472.1;
CC PIR; A03194; FDFLAW.
CC PIR; A05161; A05161.
CC PIR; S02326; S02326.
CC PIR; JS0704; JS0704.
CC PDB; 1ATF; 15-OCT-94.
CC PDB; 1WEA; 03-JUN-95.
CC PDB; 1WFB; 03-JUN-95.
CC InterPro; IPR000104; Antifreeze_1.
CC PRINTS; PR00308; ANTIFREEZE1.

```

KW Antifreeze protein; Repeat; Multigene family; Signal; 3D-structure.
FT SIGNAL 1 21
FT PROPEP 22 44 REMOVED BY A DIPEPTIDYLPEPTIDASE
FT CHAIN 45 82 (PROBABLE).
FT VARIANT 36 36 ANTIFREEZE PROTEIN A/B.
FT VARIANT 70 70 A -> V.
FT CONFLICT 24 24 A -> D (IN PROTEIN B).
FT HELIX 46 80 S -> R (IN REF. 2).
SQ SEQUENCE 82 AA; 7711 MW; C2AE7B74C0D46CC1 CRC64;
Query Match 3.2%; Score 7; DB 1; Length 82;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 120 AAAAPAA 126
Db 35 AAAAPAA 41
Search completed: July 7, 2003, 15:09:13
Job time : 24 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2003, 15:06:02 ; Search time 86 seconds
(without alignments)
524.701 Million cell updates/sec

Title: US-09-923-236-2
Perfect score: 219
Sequence: 1 MKLLWLCIVCVAFARRRF.....TAKPAPEHPSPSLEQANQ 219

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 671580 seqs, 206047115 residues

Word size : 0

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

- SPTREMBL_21.*
- 1: sp_archaea.*
 - 2: sp_bacteria.*
 - 3: sp_fungi.*
 - 4: sp_human.*
 - 5: sp_invertebrate.*
 - 6: sp_mammal.*
 - 7: sp_mhc.*
 - 8: sp_organelle.*
 - 9: sp_phage.*
 - 10: sp_plant.*
 - 11: sp_rodent.*
 - 12: sp_virus.*
 - 13: sp_vertebrate.*
 - 14: sp_unclassified.*
 - 15: sp_rvirus.*
 - 16: sp_bacteriap.*
 - 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	4.1	75	5 Q9VBL6	Q9vbl6 drosophila
2	9	4.1	344	11 Q9D5J3	Q9d5j3 mus musculus
3	9	4.1	728	5 Q9V777	Q9v777 drosophila
4	9	4.1	841	5 Q9NEA5	Q9nea5 leishmania
5	8	3.7	139	16 Q8Y2D2	Q8y2d2 ralstonia s
6	8	3.7	141	16 Q8Y2R3	Q8y2r3 ralstonia s
7	8	3.7	145	10 Q9F516	Q9f516 arabidopsis
8	8	3.7	145	10 Q9FFC0	Q9fcf7 arabidopsis
9	8	3.7	193	10 Q9C7F7	Q9c7f7 arabidopsis
10	8	3.7	209	15 Q9DQ08	Q9dqu8 human immun
11	8	3.7	228	16 Q9A4A2	Q9a4a2 caulobacter
12	8	3.7	231	2 Q93ND2	Q93nd2 myxococcus
13	8	3.7	259	2 Q91UQ4	Q91uq4 rhizobium m
14	8	3.7	261	11 Q9CSG8	Q9csg8 mus musculus
15	8	3.7	262	10 Q8W2Q2	Q8w2q2 oryza sativ
16	8	3.7	265	11 Q9CSK8	Q9csk8 mus musculus

17	8	3.7	268	16 P95192	P95192 mycobacteri
18	8	3.7	272	10 Q8S2S9	Q8s2s9 thellungiel
19	8	3.7	276	16 Q92WH2	Q92wh2 rhizobium m
20	8	3.7	291	10 Q40478	Q40478 nicotiana t
21	8	3.7	302	5 Q9N9A5	Q9n9a5 leishmania
22	8	3.7	325	2 Q9Z9X8	Q9z9x8 frateuria s
23	8	3.7	336	16 Q9AB24	Q9ab24 caulobacter
24	8	3.7	362	11 Q9CX00	Q9cx00 mus musculu
25	8	3.7	381	2 Q86996	Q86996 acinetobact
26	8	3.7	384	16 Q9AD02	Q9ad02 streptomyce
27	8	3.7	428	10 Q94I08	Q94iq8 zea mays (m
28	8	3.7	447	8 Q9TN70	Q9tn70 tupistra al
29	8	3.7	448	8 Q33130	Q33130 streltiztia
30	8	3.7	468	8 Q9GEQ0	Q9geq0 primula cor
31	8	3.7	493	2 Q48431	Q48431 klebsiella
32	8	3.7	493	5 Q9W4P4	Q9w4p4 drosophila
33	8	3.7	493	5 Q8S2R1	Q8s2r1 drosophila
34	8	3.7	495	10 Q9XEN0	Q9xen0 oryza sativ
35	8	3.7	512	5 Q9VM14	Q9vml4 drosophila
36	8	3.7	550	16 Q9RY21	Q9ry21 deinococcus
37	8	3.7	551	3 Q42691	Q42691 alternaria
38	8	3.7	555	2 Q85737	Q85737 streptomyce
39	8	3.7	569	10 Q8W2Q4	Q8w2q4 oryza sativ
40	8	3.7	600	16 Q8XSG9	Q8xsg9 ralstonia s
41	8	3.7	611	16 Q92TK5	Q92tk5 rhizobium m
42	8	3.7	679	2 Q9ZAS9	Q9zas9 ralstonia s
43	8	3.7	715	11 Q63803	Q63803 rattus norv
44	8	3.7	789	16 Q8XX15	Q8xx15 ralstonia s
45	8	3.7	842	10 Q9LEE7	Q9lee7 zea mays (m

ALIGNMENTS

RESULT 1

ID	Q9VBL6	PRELIMINARY;	PRT;	75 AA.
AC	Q9VBL6; Q24390;			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)			
DE	Accessory gland-specific peptide 57Da precursor (Male accessory gland secretory protein 57Da).			
GN	MST57DA OR CG9074.			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_TaxID=7227;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.			
RC	STRAIN=OREGON-R;			
RX	MEDLINE=95227188; PubMed=7711745;			
RA	Simmerl E., Schaefer M., Schaefer U.;			
RT	"Structure and regulation of a gene cluster for male accessory gland transcripts in Drosophila melanogaster".			
RL	Insect Biochem. Mol. Biol. 25:127-137(1995).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=BERKELEY;			
RX	MEDLINE=20196006; PubMed=10731132;			
RA	Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,			
RA	Ananadis P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,			
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,			
RA	Sutton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,			
RA	Brannon R.C., Rogers Y.-H.C., Blazek R.G., Champagne M., Pfeiffer B.D.,			
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,			
RA	Abriel J.F., Abayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,			
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,			
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,			
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,			
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,			
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,			

```
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Liao X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarly K., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamis I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
CC -1- FUNCTION: THIS PROTEIN IS TRANSFERRED FROM MALE TO FEMALE DURING
CC MATING AND MAY AFFECT EGG-LAYING AND BEHAVIOR AFTER MATING.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- TISSUE SPECIFICITY: LUMEN FLUID OF MALE ACCESSORY GLANDS, BECOMES
CC SEMINAL FLUID.
DR EMBL: Z33647; CAA83925.1; -.
DR EMBL: AE003753; AAP56515.1; -.
DR FlyBase: FBgn0011668; Mst57Da.
KW Signal; Behavior.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 75 ACCESSORY GLAND-SPECIFIC PEPTIDE 57DA.
FT DOMAIN 39 64 ALA/PRO-RICH DOMAIN.
FT CONFLICT 39 64 MISSING (IN REF. 1).
FT CONFLICT 64 75 MISSING (IN REF. 1).
FT SEQUENCE 75 AA; 7106 MW; BCD3D70817C98E14 CRC64;
SQ
Query Match 4.1%; Score 9; DB 5; Length 75;
Best Local Similarity 100.0%; Pred. No. 0.65;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 120 AAAAPAAP 128
DB 45 AAAAPAAP 53
RESULT 2
Q9D5J3 ID Q9D5J3 PRELIMINARY; PRT; 344 AA.
AC Q9D5J3
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE 4930432K09Rik protein.
GN 4930432K09Rik
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TESTIS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa K., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
```

```
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gari B.D.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-P.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Hayashaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL: AK015291; BAB29782.1; -.
DR MGD: MGI:1921029; 4930432K09Rik.
DR InterPro: IPR002965; P-rich_extensn.
DR PRINTS: PR01217; PRICEXTENSN.
DR SEQUENCE 344 AA; 36294 MW; D65EAD71CE5802AC CRC64;
SQ
Query Match 4.1%; Score 9; DB 11; Length 344;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKLLWACI 9
DB 1 MKLLWACI 9
RESULT 3
Q9VY77 ID Q9VY77 PRELIMINARY; PRT; 728 AA.
AC Q9VY77
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CG11063 protein.
GN CG11063
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abrell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Liao X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarly K., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamis I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
CC -1- FUNCTION: THIS PROTEIN IS TRANSFERRED FROM MALE TO FEMALE DURING
CC MATING AND MAY AFFECT EGG-LAYING AND BEHAVIOR AFTER MATING.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- TISSUE SPECIFICITY: LUMEN FLUID OF MALE ACCESSORY GLANDS, BECOMES
CC SEMINAL FLUID.
DR EMBL: Z33647; CAA83925.1; -.
DR EMBL: AE003753; AAP56515.1; -.
DR FlyBase: FBgn0011668; Mst57Da.
KW Signal; Behavior.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 75 ACCESSORY GLAND-SPECIFIC PEPTIDE 57DA.
FT DOMAIN 39 64 ALA/PRO-RICH DOMAIN.
FT CONFLICT 39 64 MISSING (IN REF. 1).
FT CONFLICT 64 75 MISSING (IN REF. 1).
FT SEQUENCE 75 AA; 7106 MW; BCD3D70817C98E14 CRC64;
SQ
Query Match 4.1%; Score 9; DB 11; Length 344;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKLLWACI 9
DB 1 MKLLWACI 9
RESULT 3
Q9VY77 ID Q9VY77 PRELIMINARY; PRT; 728 AA.
AC Q9VY77
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CG11063 protein.
GN CG11063
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abrell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Liao X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarly K., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
```

RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
CC -1- SIMILARITY: CONTAINS 3 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC
CC IONS.
DR EMBL; AE003493; AAF48328.1; -.
DR HSSP; P04006; 1IML.
DR FlyBase; FBgn0030530; CG11063.
DR InterPro; IPR001781; LIM.
DR Pfam; PF00412; LIM; 3.
DR PRODOM; PD000094; LIM; 3.
DR SMART; SM00132; LIM; 3.
DR PROSITE; PS00478; LIM_DOMAIN.1; 2.
DR PROSITE; PS0023; LIM_DOMAIN.2; 3.
KW LIM domain; Metal-binding; Zinc.
SQ SEQUENCE 728 AA; 77961 MW; 86F9CAFF840E4541 CRC64;
Query Match 4.1%; Score 9; DB 5; Length 728;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 119 SAAAPAAP 127
Db 418 SAAAPAAP 426
RESULT 4
Q9NEA5 ID Q9NEA5 PRELIMINARY; PRT; 841 AA.
AC Q9NEA5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Possible inner dynein arm.
GN L5515.03.
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OC NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FRIEDLIN;
RA Masuy D., Purnelle B., Goffeau A., Ivens A.C., Quail M.,
RA Rajandream M.A., Barrell B.G.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FRIEDLIN;
RX MEDLINE=98146435; PubMed=9477341;
RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
RA Smith D.F.;
RT "A physical map of the Leishmania major Friedlin genome.";
RL Genome Res. 8:135-145(1998).
CC -1- SIMILARITY: CONTAINS 4 WD REPEATS (TRP-ASP DOMAINS).
DR EMBL; AL161414; CAB77677.1; -.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 4.
DR SMART; SM00320; WD40; 4.
DR PROSITE; PS00678; WD_REPEATS.1; UNKNOWN_2.
DR PROSITE; PS0294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 841 AA; 93125 MW; B41B72F1B8821226 CRC64;
Query Match 4.1%; Score 9; DB 5; Length 841;

Best Local Similarity 100.0%; Pred. No. 5,7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 119 SAAAPAAP 127
Db 40 SAAAPAAP 48
RESULT 5
Q8Y2D2 ID Q8Y2D2 PRELIMINARY; PRT; 139 AA.
AC Q8Y2D2;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical signal peptide protein RSC0404.
GN RSC0404 OR RS03372.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC Ralstonia.
OC NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GM11000;
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chandler M., Choise N., Claudel-Renard C., Cunnac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA Sigulier P., Thebault P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002).
DR EMBL; AL646059; CAD13932.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 139 AA; 14153 MW; A31350B6FD629039 CRC64;
Query Match 3.7%; Score 8; DB 16; Length 139;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 120 AAAAPAAP 127
Db 55 AAAAPAAP 62
RESULT 6
Q8Y2R3 ID Q8Y2R3 PRELIMINARY; PRT; 141 AA.
AC Q8Y2R3;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical protein RSC0272.
GN RSC0272 OR RS03239.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC Ralstonia.
OC NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GM11000;
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chandler M., Choise N., Claudel-Renard C., Cunnac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA Sigulier P., Thebault P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002).
DR EMBL; AL646058; CAD13800.1; -.
KW InterPro; IPR001763; Rhodanese-like.

```

DR SMART; SM00450; RHOD; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 141 AA; 15454 MW; 878FE1840A98B4E CRC64;

Query Match      3.7%  Score 8; DB 16; Length 141;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 149 AAGAPVAA 156
Db 126 AAGAPVAA 133

RESULT 7
Q96516 ID Q96516 PRELIMINARY; PRT; 145 AA.
AC Q96516;
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Histone H2B.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ECOTYPE COLUMBIA; TISSUE=CELL SUSPENSION CULTURE;
RA Phillips G.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT: THE NUCLEOSOME IS AN OCTAMER CONTAINING TWO MOLECULES OF
CC H2A, H2B, H3, AND H4; WHICH WRAP APPROXIMATELY 146 BP OF DNA (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; Y07745; CAA69025.1; -.
DR InterPro; IPR004822; Histone_core.
DR Pfam; PF00125; histone; 1.
DR PRINTS; PR00621; HISTONEH2B.
DR ProDom; PD000497; Histone_H2B; 1.
DR SMART; SM00427; H2B; 1.
DR PROSITE; PS00357; HISTONE_H2B; 1.
KW Chromosomal protein; DNA-binding; Nuclear protein; Nucleosome core.
SQ SEQUENCE 145 AA; 15733 MW; 6C8EB8B18390F4686 CRC64;

Query Match      3.7%  Score 8; DB 10; Length 145;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 130 AAEPAAAA 137
Db 19 AAEPAAAA 26

RESULT 8
Q9FFC0 ID Q9FFC0 PRELIMINARY; PRT; 145 AA.
AC Q9FFC0;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Histone H2B.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RA White O.; Alonso J.; Altafi H.; Araujo R.; Bowman C.L.; Brooks S.Y.;
RA Buehler E.; Chan A.; Chao Q.; Chen H.; Cheuk R.F.; Chin C.W.;
RA Chung M.K.; Conn L.; Conway A.B.; Conway A.R.; Creasy T.H.; Dewar K.;
RA Dunn P.; Etgu P.; Feldblyum T.V.; Feng J.-D.; Fong B.; Fujii C.Y.;
RA Gill J.E.; Goldsmith A.D.; Haas B.; Hansen N.F.; Hughes B.; Huizar L.;
RA Hunter J.L.; Jenkins J.; Johnson-Hopson C.; Khan S.; Khaykin E.;
RA Kim C.J.; Koo H.L.; Kremenetskaia I.; Kurtz D.B.; Kwan A.; Lam B.;
RA Langin-Hooper S.; Lee A.; Lee J.M.; Lenz C.A.; Li J.H.; Li Y.-P.;
RA Lin X.; Liu S.X.; Liu Z.A.; Luros J.S.; Maiti R.; Marziani A.;
RA Militischer J.; Miranda M.; Nguyen M.; Nierman W.C.; Osborne B.I.;
RA Pai G.; Peterson J.; Pham P.K.; Rizzo M.; Rooney T.; Rowley D.;
RA Sakano H.; Salzberg S.L.; Schwartz J.R.; Shinn P.; Southwick A.M.;
RA Sun H.; Tallon L.J.; Tamburiga G.; Toriumi M.J.; Town C.D.;
RA Utterback T.; Van Aken S.; Vaysberg M.; Vysotskaia V.S.; Walker M.;
RA Wu D.; Yu G.; Fraser C.M.; Venter J.C.; Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
thaliana."
RL Nature 408:816-820(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Southwick A.; Karlin-Neumann G.; Nguyen M.; Lam B.; Miranda M.;
RA Palm C.J.; Bowser L.; Jones T.; Banh J.; Carninci P.; Chen H.;
RA Cheuk R.; Chung M.K.; Hayashizaki Y.; Ishida J.; Kamiya A.; Kawai J.;

```

```

RA Sato S.; Kotani H.; Nakamura Y.; Kaneko T.; Asamizu E.; Fukami M.;
RA Miyajima N.; Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. I. Sequence
RT features of the 1.6 Mb regions covered by twenty physically assigned
RT PI clones."
RL DNA Res. 4:215-230(1997).
CC -1- SUBUNIT: THE NUCLEOSOME IS AN OCTAMER CONTAINING TWO MOLECULES OF
CC H2A, H2B, H3, AND H4; WHICH WRAP APPROXIMATELY 146 BP OF DNA (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE HISTONE H2B FAMILY.
DR EMBL; AB005243; BAB10609.1; -.
DR InterPro; IPR004822; Histone_core.
DR Pfam; PF00125; histone; 1.
DR PRINTS; PR00621; HISTONEH2B.
DR ProDom; PD000497; Histone_H2B; 1.
DR SMART; SM00427; H2B; 1.
DR PROSITE; PS00357; HISTONE_H2B; 1.
KW Chromosomal protein; DNA-binding; Nuclear protein; Nucleosome core.
SQ SEQUENCE 145 AA; 15732 MW; CC8421B559D42B30 CRC64;

Query Match      3.7%  Score 8; DB 10; Length 145;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 130 AAEPAAAA 137
Db 19 AAEPAAAA 26

RESULT 9
Q9C7F7 ID Q9C7F7 PRELIMINARY; PRT; 193 AA.
AC Q9C7F7;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Lipid transfer protein, putative.
GN FLK9.6 OR ATIG27950.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A.; Ecker J.R.; Palm C.J.; Federspiel N.A.; Kaul S.; S.Y.;
RA White O.; Alonso J.; Altafi H.; Araujo R.; Bowman C.L.; Brooks S.Y.;
RA Buehler E.; Chan A.; Chao Q.; Chen H.; Cheuk R.F.; Chin C.W.;
RA Chung M.K.; Conn L.; Conway A.B.; Conway A.R.; Creasy T.H.; Dewar K.;
RA Dunn P.; Etgu P.; Feldblyum T.V.; Feng J.-D.; Fong B.; Fujii C.Y.;
RA Gill J.E.; Goldsmith A.D.; Haas B.; Hansen N.F.; Hughes B.; Huizar L.;
RA Hunter J.L.; Jenkins J.; Johnson-Hopson C.; Khan S.; Khaykin E.;
RA Kim C.J.; Koo H.L.; Kremenetskaia I.; Kurtz D.B.; Kwan A.; Lam B.;
RA Langin-Hooper S.; Lee A.; Lee J.M.; Lenz C.A.; Li J.H.; Li Y.-P.;
RA Lin X.; Liu S.X.; Liu Z.A.; Luros J.S.; Maiti R.; Marziani A.;
RA Militischer J.; Miranda M.; Nguyen M.; Nierman W.C.; Osborne B.I.;
RA Pai G.; Peterson J.; Pham P.K.; Rizzo M.; Rooney T.; Rowley D.;
RA Sakano H.; Salzberg S.L.; Schwartz J.R.; Shinn P.; Southwick A.M.;
RA Sun H.; Tallon L.J.; Tamburiga G.; Toriumi M.J.; Town C.D.;
RA Utterback T.; Van Aken S.; Vaysberg M.; Vysotskaia V.S.; Walker M.;
RA Wu D.; Yu G.; Fraser C.M.; Venter J.C.; Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
thaliana."
RL Nature 408:816-820(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Southwick A.; Karlin-Neumann G.; Nguyen M.; Lam B.; Miranda M.;
RA Palm C.J.; Bowser L.; Jones T.; Banh J.; Carninci P.; Chen H.;
RA Cheuk R.; Chung M.K.; Hayashizaki Y.; Ishida J.; Kamiya A.; Kawai J.;

```



```

RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.,
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC069471; AAG51485.1; -.
DR EMBL; AY092956; AAM12955.1; -.
DR InterPro; IPR003612; AAI.
DR InterPro; IPR001768; TRY/amy1_inhtr.
DR Pfam; PF00234; TRY_alpha_aml1; 1.
DR SMART; SM00499; AAI_1.
SQ SEQUENCE 193 AA; 19759 MW; D54B38B12FFB6610 CRC64;

Query Match 3.7%; Score 8; DB 10; Length 193;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 AAAAPAAP 127
Db 20 AAAAPAAP 27

RESULT 10
Q9DQ08 PRELIMINARY; PRT; 209 AA.
AC Q9DQ08;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DE Negative factor (F-protein) (27 kDa protein).
GN NEF.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=01U09;
RX MEDLINE=21002575; PubMed=11118071;
RA Geffin R., Wolf D., Muller R., Hill M.D., Stellwag E., Freitag M.,
RA Sassi G., Scott G.B., Baur A.S.;
RT "Functional and structural defects in HIV-1 nef genes derived from
RT pediatric long-term survivors";
RL AIDS Res. Hum. Retroviruses 16:1855-1868(2000).
CC -1- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4 (T4) ANTIGEN.
DR EMBL; AF252901; AAG34573.1; -.
DR HSSP; P03406; 1EFN.
DR InterPro; IPR001558; HIV_Nef.
DR Pfam; PF00469; F-protein; 1.
DR ProDom; PD000031; HIV_Nef; 1.
DR AIDS; GTP-binding; Myristate.
KW AIDS; GTP-binding; Myristate.
SQ SEQUENCE 209 AA; 23656 MW; 6B858BAD98C79519 CRC64;

Query Match 3.7%; Score 8; DB 15; Length 209;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 AEPAAAAP 138
Db 23 AEPAAAAP 30

RESULT 11
Q9A4A2 PRELIMINARY; PRT; 228 AA.
AC Q9A4A2;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Cytochrome c, membrane-bound.
GN CC2935.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
OC Caulobacter.

```

```

OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647;
RA Eisen J., Heideberg J.F., Alley M.R.K., Ohta N., Maddock J.R., Ely B.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Haft D.H.,
RA Deboy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Uterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL; AE005958; AAK24897.1; -.
DR HSSP; P81459; 1I54.
DR TIGR; CC2935; -.
DR InterPro; IPR00104; Antifreeze_1.
DR InterPro; IPR000345; CytC_heme_bind.
DR InterPro; IPR003088; Cyt_C1.
DR InterPro; IPR002327; Cyt_C1AB.
DR Pfam; PF00034; cytochrome_c; 1.
DR PRINTS; PR00308; ANTIFREEZE1.
DR PRINTS; PR00604; CYTCHRMCEIAB.
DR ProDom; PD000375; Cyt_C1AB; 1.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 228 AA; 23023 MW; 5AE05C7DD5D05F82 CRC64;

Query Match 3.7%; Score 8; DB 16; Length 228;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 AAAAPAAP 127
Db 219 AAAAPAAP 226

RESULT 12
Q93ND2 PRELIMINARY; PRT; 231 AA.
AC Q93ND2;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Hypothetical 24.5 kDa protein.
OS Myxococcus xanthus.
OC Bacteria; Proteobacteria; delta subdivision; Myxobacteria;
OC Myxococcales; Cystobacterineae; Myxococcaceae; Myxococcus.
OX NCBI_TaxID=34;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DZFl;
RA Nariya H., Inouye S.;
RT "Identification of serine/threonine kinase associate proteins in M.
RT xanthus by yeast two-hybrid system.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF377338; AAK64438.1; -.
KW Hypothetical protein.
SQ SEQUENCE 231 AA; 24488 MW; 8A2AA9FF8D789536 CRC64;

Query Match 3.7%; Score 8; DB 2; Length 231;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 149 AAGAPVAA 156
Db 39 AAGAPVAA 46

RESULT 13
Q91UQ4 PRELIMINARY; PRT; 259 AA.
ID Q91UQ4

```

```
AC Q9IU04;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE VirB1 protein.
GN VIRB1.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Plasmid pSB102.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC TRANSPOSON-MERCURY RESISTANCE TRANSPOSON TN5718;
RA Schneider S., Keller M., Droge M., Lanka E., Puehler A.,
RA Selbitschka W.;
RT "The genetic organization and evolution of the broad-host-range
RT mercury resistance plasmid pSB102 isolated from a microbial population
RT residing in the rhizosphere of alfalfa.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ304453; CAC79183.1; -.
DR InterPro; IPR001179; FKBP_PPIase.
DR InterPro; IPR000189; SLT_domain.
DR Pfam; PF01464; SLT; 1.
DR PROSITE; PS00453; FKBP_PPIASE_1; UNKNOWN_1.
KW Plasmid.
SQ SEQUENCE 259 AA; 27380 MW; 09648108175E9A7B CRC64;

Query Match 3.7%; Score 8; DB 2; Length 259;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 152 APVAAEPA 159
Db 193 APVAAEPA 200
|||||||

RESULT 14
Q9CSG8
ID Q9CSG8 PRELIMINARY; PRT; 261 AA.
AC Q9CSG8;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE 1810019E15RIK protein (Fragment).
GN 1810019E15RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-C57BL/6J; TISSUE=EMBRYO;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gajobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Giesi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
```

```
RL Nature 409:685-690(2001).
DR EMBL: AK012877; BAB8530.1; -.
DR MGD; MGI:1914254; 1810019E15RIK.
DR InterPro; IPR003034; SAP.
DR Pfam; PF02037; SAP; 1.
DR SMART; SM00513; SAP; 1.
FT NON_TER 261 261
SQ SEQUENCE 261 AA; 29664 MW; D5CA7CDC443FAA9E CRC64;

Query Match 3.7%; Score 8; DB 11; Length 261;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 119 SAAAPAA 126
Db 2 SAAAPAA 9
|||||||

RESULT 15
Q8W2Q2
ID Q8W2Q2 PRELIMINARY; PRT; 262 AA.
AC Q8W2Q2;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Putative polyprotein.
GN OSJNBA0028C16.20.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Wing R.A., Yu Y., Soderlund C., Chen M., Kim H.-R., Rambo T.,
RA Sasaki C., Henry D., Oates R., Simmons J.;
RT "Rice Genomic Sequence.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC098565; AAL69438.1; -.
KW Polyprotein.
SQ SEQUENCE 262 AA; 27461 MW; C7BA920CF2486429 CRC64;

Query Match 3.7%; Score 8; DB 10; Length 262;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 121 AAAPAAP 128
Db 148 AAAPAAP 155
|||||||

Search completed: July 7, 2003, 15:12:07
Job time : 88 secs
```